A PROPOSED FRAMEWORK FOR FORENSIC IMAGE ENHANCEMENT

by

SPENCER AGUILA LEDESMA

B.A., University of Northern Colorado, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Recording Arts

2015

This thesis for the Master of Science degree by

Spencer Aguila Ledesma

has been approved for the

Recording Arts Program

by

Catalin Grigoras

Jeffrey Smith

Scott Burgess

November 20, 2015

Ledesma, Spencer Aguila (M.S. Recording Arts)

A Proposed Framework for Forensic Image Enhancement

Thesis directed by Associate Professor Catalin Grigoras

ABSTRACT

Digital images and videos used in the investigation of a crime often undergo several concurrent enhancement operations for improved analysis by humans or automated systems. When applying multiple image processing techniques to an image, the order and method in which processes are applied can have a profound impact on the result. However, the effect that one enhancement algorithm will have when applied in conjunction with another is not always obvious. When applied incorrectly, at best, there will be a negative impact to the amount of information that can be extracted from an image. At worst, the information contained in a processed image could be misrepresented. This thesis proposes a tool independent workflow for forensic image enhancement with a strong emphasis on an order of operations that maximizes the efficacy of each enhancement technique while observing the responsibilities and best practices of the forensic science community. This work will be useful for developing an understanding of common image enhancement techniques, understanding how these techniques relate to forensic science, and aiding in the creation of quality assurance standards for forensic image enhancement. Chapter 1 gives an introduction to image enhancement and discusses its role in forensic science and litigation. Chapter 2 summarizes the digital image creation process and its relationship to the human visual system. Chapter 3 reviews the most commonly used image enhancement techniques, including their theoretical background, strengths, and limitations. Chapter 4 introduces a framework for image enhancement and the rationale behind it through a series of practical examples.

The form and content of this abstract are approved. I recommend its publication.

Approved: Catalin Grigoras

DEDICATION

I dedicate this work to my family: Esther, Abel, Jose, and Dora, whom have always shown me unconditional support and love. Everything good that I have or will accomplish in life is because of you.

ACKNOWLEDGEMENTS

I would like to thank my professors, Dr. Catalin Grigoras and Jeff Smith, for guiding me in my studies. This work would not have been possible without their abundant knowledge, support, and patience. Their passion for forensic science along with their wholehearted dedication to me and my class is a constant inspiration.

Furthermore, I'd like to thank Dr. Leonid Rudin for the privilege of allowing me access to a wide range of experiences, forensic tools, and his years of wisdom to aid my development and understanding in the field of forensics and image processing.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

FIGURE

LIST OF TABLES

CHAPTER I

INTRODUCTION

Image enhancement is an accepted practice in the field of digital & multimedia forensics and heavily relied upon in many forensic applications such as crime scene reconstruction, photogrammetry, questioned documents, and biometric analysis including facial and finger print identification [1]. It is not uncommon for images used in these applications to undergo several concurrent image processing operations. When multiple processing operations are applied to an image, it is significant to note that like a falling stack of dominos, each operation that an image undergoes will have an effect on any future processing of the image. Even when applying the exact same enhancement techniques, at the same exact settings, to the same image, applying enhancements in the *incorrect* order can lead to an overall loss in image fidelity or the creation of features that are non-existent in the original image data, including artifacts like image noise or false edges.

Visual components of digital images are in principle a matrix of numerical values. Image processing operations use algorithms to manipulate these numerical values mathematically. Since these algorithms operate in predefined ways, it is possible to predict their behavior. By studying the underlying processes of enhancement algorithms it is therefore possible to predict how they react in relation to different image properties and thereby establish an ideal order for their application.

It goes without saying that the great stakes involved in a forensic investigation demand that ideal methods be followed whenever possible. However, defining an ideal can be challenging as ideals are ultimately subjective and only become meaningful in relation to an ultimate goal. Therefore it is necessary to understand what is meant by the word “ideal” under the purview of forensic science. After all, what may be an ideal use of image processing for the sake of film special effects, for example, is not at all ideal for forensic applications. The goals of forensic image enhancement will be discussed in detail throughout this chapter but for now, the following simplified definition will be adequate: The goal of forensic image enhancement is to improve the visual appearance of an image in order for it to become more useful in scientific investigations and legal matters[15][17]. Under this definition, any order of operations for image processing that yields more reliable information from an image without misrepresenting the imaged content would be considered a more ideal method than another.

Traditionally, the order in which image enhancements are applied has not been universally agreed upon by the media forensics community. Since the needs of every enhancement case are unique, requiring different combinations of image processing operations and settings, establishing an order of operations for image enhancement has often been seen as impossible, or at the least, impractical. This thesis advocates that an order of operations for forensic digital image enhancement is both possible and readily applicable to many common enhancement use cases.

In defense of this thesis it will be necessary to explore how image processing algorithms can interact with one another. Though image enhancement is an established subject heavily studied in scientific publications, literature on this topic is frequently concerned with individual techniques and has not often been explored holistically from the perspective of forensic science. Therefore it is necessary to bridge the gap between image processing theory and the principles of forensics. As the foundation for this discussion, this thesis gives an overview of the current state of forensics, guiding principles in digital multimedia forensics, the human visual system, the digital image creation pipeline, and common digital image enhancement techniques. This foundation will be used for the proposal of a workflow for image enhancement that maximizes the efficacy of each enhancement, minimizes the creation of unwanted image artifacts, and conforms to forensic principles and best practices.

Scope

First and foremost, this thesis strives to be beneficial to the digital & multimedia forensics and law enforcement communities for the purposes of quality assurance such as the development of standard operating procedures, best practices, and training materials. Nonetheless, this work aims to provide sufficient background information in order to be accessible to those interested in this subject who are outside of these fields, including legal professionals and newcomers to forensics. Though much of the content presented in this paper relating to case law and best practices are written from the perspective of forensics as it is practiced in the United States of America, many of these practices have been developed through a consensus of international scientific and legal communities. The information presented in this paper on these subjects is therefore applicable to the broad international audience. While this paper takes a comprehensive approach to give the reader an understanding of the current climate of forensic image enhancement, it is important to note that the level of knowledge necessary to become qualified as a forensic image enhancement practitioner is variable and most certainly outside of this paper's scope. Necessary qualifications for forensic image enhancement practitioners are contingent on the types of enhancements used, their intended use, legal jurisdiction, and the standards and bylaws of several overseeing bodies, as will be discussed further on. Readers interested in an advanced understanding of the mathematics involved in image processing and other topics beyond the scope of this work are encouraged to seek out the great body of literature on this subject including the publications referenced within this paper.

The image enhancements presented in this document are constrained to *digital* images, including those found in digital videos, and the *visual* data contained within. This paper does not present or discuss image authentication procedures or techniques. *Enhancement* and analysis of non-visual data, including metadata and algorithms for generating statistical representations of data for authentication or analysis purposes are beyond the scope of this paper.

The approach taken in this work is tool agnostic, that is, without endorsement of any specific tool set. While common tools may be referenced within this work or used for its creation, readers will not find a step by step instruction on the use of specific enhancement filters within a specific software program. This work does not in any way claim to be a comprehensive list of all currently available digital image enhancement techniques found in all commercial programs. Only the most commonly available and used image enhancement techniques with suitability to forensics are discussed. The information contained in this paper regarding the underlying processes of enhancement techniques is presented as a theoretical foundation from the perspective of a forensic image analyst and/or forensic image enhancement practitioner. This information will not delve into details of algorithms from specific programs. It would be impractical to discuss specific algorithms as they are often tool specific, well-guarded by their respective developers, and subject to frequent revision [2].

Public perception

As the proliferation of inexpensive electronics continues to grow with no end in sight, so has the spread of digital cameras. The latest estimates from the Consumer Electronics Association predict that 85% of American households have some kind of digital camera and this number is rising every year [3]. In addition to handheld digital cameras, Digital Video Recorder (DVR) based surveillance systems, police body cameras, camera phones, etc., are all contributing to a rise in digital image and video evidence. Though the quality of digital camera technology is constantly improving, digital images used in forensic investigations are often less than ideal. Problems such as poor resolution, poor contrast, inaccurate color reproduction, blur, noise, and user error are typical problems encountered with digital image evidence. Image enhancement can be used to correct or improve the effect of many of these problems so that image evidence can better serve the legal system, but often images may be too degraded for enhancements to be effective.

Unfortunately, there is much confusion and unrealistic expectations regarding which types of enhancements are possible and which are not. This confusion is in no small part due to portrayals of forensics in television crime dramas which often tout fiction over fact in the name of good television. Fictional crime related TV has been criticized in recent years for inaccurate depictions of forensic science, where image enhancement and other forensic techniques are wielded as some kind of "high-tech magic" [4]. Some estimate that up to 40% of forensic techniques used in popular television do not exist [5]. This depiction of forensics on TV has been noted to cause regular viewers of these shows to perceive forensic science as some type of unfaltering super science, creating unrealistically high expectations of forensic science in US court rooms, and changing the outcome of legal rulings [6]. This so called "CSI Effect", aptly named after the popular television crime series, is concerning as figures indicate that 100 million Americans, roughly one third of the US population, watch crime related television regularly [7].

To add to this confusion, the accessibility of image processing tools to the general public has made it difficult to differentiate forensically acceptable and inacceptable enhancement practices. The typical use of image enhancement outside of a forensic context deals with making an image aesthetically pleasing, while forensic image enhancement is generally a restorative practice used to make an image more useful to an investigation. The irony is that despite these two disparate goals, many of the image processing tools used by forensic image enhancement practitioners are the same tools being used by the public for aesthetic purposes. The steady beat of Moore's Law has resulted in public accessibility of image processing techniques that could once only be found in specialized hardware and professional level software suites to be freely applied with an inexpensive computer or smartphone. Professor of Engineering, Jeff Bokor at the UC Berkeley estimates that the Apple iPhone 6, last year's model as of this writing, is roughly 1,000,000 times faster than an IBM computer from 1975 which took an entire room [8]! This state of affairs is blurring public distinctions between forensic and non-forensic image processing.

Forensic science

In order to separate science fact from fiction it is necessary to first have a clear understanding of what forensic science is. Science is a systematic approach for understanding how we know what we know. Science does not need a laboratory and white lab coats; it only requires that one is willing to test a hypothesis, evaluate evidence, and follow the evidence to its logical conclusions. The eloquent words of Carl Sagan say it best when in his final public interview he declared "Science is more than a body of knowledge. It is a way of thinking; a way of skeptically interrogating the universe with a fine understanding of human fallibility" [9]. Science allows us to humble ourselves to entertain the possibility that we can be wrong about our preconceptions, intuitions, and deeply held beliefs. Rather than attempting to prove what we already believe to be true, or what we want to be true, the scientific method demands that we attempt to disprove ourselves. Hypotheses that cannot be disproven are retested mercilessly and results are peer reviewed and tested again, until findings can be deemed reliable enough to join the lofty ranks of scientific theory. This is not to say that science can give us absolute truth. Science can only lead to reasonable expectations given the available evidence. All scientific theories, no matter how established they may be, are subject to revision in accordance with the best available evidence.

Forensics is the application of science toward matters relevant to the legal system. Forensic scientists *interrogate the universe* to piece together the most reasonable interpretation of the available evidence after the occurrence of a crime. The findings of a forensic investigation are not ends unto themselves. For any findings to be of use to the legal system, they must be presented, and in a manner that can be easily understood. The presentation of the fruits of scientific inquiry is the ethical responsibility of not only forensic scientists but all science as a whole. Again, Sagan eloquently encapsulated this sentiment when he wrote the following: "Not explaining science seems to me perverse. When you're in love, you want to tell the world" [10]. Forensic scientists serve equally important roles as both investigators and teachers. It is the role of a forensic scientist to disavow the public of misconceptions and disseminate accurate, scientifically founded knowledge. Thus forensic science follows four basic stages for processing evidence: retrieval of evidence, analysis, interpretation, and presentation of findings.

All stages of a forensic investigation operate on the acknowledgement of Locard's Exchange Principle, the concept that every action leaves a trace [11]. Locard's exchange principle applies not only to the evidence left by perpetrators or victims of a crime but also extends to the actions of forensic analysts who themselves will leave traces of their involvement with evidence and crime scenes. As any changes to the way that evidence is acquired, handled, analyzed, and presented can have direct consequences on the outcome of a legal ruling, forensic science demands the careful, unbiased handling of evidence, and presentation of findings by knowledgeable experts.

Forensic image enhancement

An image is representation of a person or thing, drawn, painted, photographed, etc. [12]. In digital photography, these representations are created by a matrix of numerical values which are interpreted as units of color by a monitor or printer [12]. By altering the numerical values that a digital image consists of, the image may be enhanced, modified, or destroyed.

Forensic image enhancement falls under the oversight of Digital & Multimedia Forensics, often referred to simply as Media Forensics. Multimedia is defined as analog or digital media, including, but not limited to, film, tape, magnetic and optical media, and/or the information contained therein [12]. Media forensics oversees the acquisition, preservation, analysis, and presentation of analog and digital audio, image, and video media evidence. In accordance with the goals of forensics as a whole, the guiding principal of media forensics is to maintain the integrity and provenance of media upon seizure, and throughout the analysis and handling process [13]. This translates to ensuring that evidence is not contaminated, lost, changed, or destroyed.

Because all image enhancements in some way modify the data of a digital image and its appearance, this may lead some to believe that image enhancement is in direct contradiction to the principles of media forensics. This contradiction is an illusion. As we will see, what determines if an enhancement is forensically acceptable is dependent on the purpose for the enhancement, the methodology used throughout processing of image data, and the law.

Enhancement or manipulation?

The term image enhancement is often used to describe varying image processes that improve the visual appearance of an image. While many image processing tools exist which can make an image look "better", a clear distinction must be made between what is a forensically sound enhancement and what is best described as image manipulation. In a forensic context, image enhancement does not serve to make images aesthetically pleasing or to generate a *desired* image. Rather, it serves the purpose of revealing preexistent image information to the human eye [15] and to give a fair and accurate representation of an imaged scene.

In his seminal book, "Fundamentals of Image Processing", Anil Jain writes that the goal of image enhancement is to "accentuate certain image features for subsequent analysis or for image display" [2]. The use of the word "accentuate" is the key to this definition. It implies that image data is not created by enhancement, but rather it is made to stand out in a way that makes it more meaningful. With this definition in play, it can be argued that "enhancements" typical of fashion magazines such as skin smoothing and colorization are not true enhancements after all. Skin smoothing removes information of skin texture and image colorization introduces new color information into an image rather than extracting color information from existing image data. In contrast, enhancements such as enlargement of an image can be performed in such a way that no scene information is added or removed, making it a technique that is a better representative of Jain's definition.

It would seem that the removal or addition of non-existing data to an image goes against the principles of forensics, making this type of "enhancement" an exercise in art (or deceit) rather than forensics. However, there is a caveat. Sometimes, the removal or addition of information to an image *can* serve a purpose in matters of law. Modifying the color of an image may help make an object more discernable or true to the colors of the imaged scene. Removing image data can also accentuate image content by eliminating content that would otherwise distract. Furthermore, the removal of image information can be used to protect the identity of a victim (Fig. 1.1) or to remove information irrelevant to a case.

Figure 1.1: A mosaic filter.
This type of mosaic filter is commonly used to protect the identity of a victim.

Some image processing tools are easy to prohibit from all forensic use because they create information rather than reveal it, while others cannot be so easily dismissed. An image processing tools like the “clone tool” (Fig. 1.2), a tool that lets one copy and paste information from one part of an image to another, has no place in forensics as it is impossible to use without creating new data and giving a false representation of a photographed scene. Yet a tool like the crop tool can be used legitimately or for deception. What we can take from all this is that what defines an image processing operation as forensically sound is not dependent on the tool; it’s how the tool is used. Image enhancements are distinct from manipulations of images so long as they do not misrepresent the content of an image and forensic principles are followed throughout all stages of an investigation.

Figure 1.2: Image manipulation. [14]
These images of an Iranian missile test made headlines in 2008 when it was made apparent that images of the missile launch where edited to exaggerate the capabilities of Iran. Above is the original image. Below is the image altered to appear like an extra missile has launched. This is something that can be easily done with the clone tool.

Legal considerations

Photography has a long history in courts of law. Seeing as an average person will only remember 65% of what he or she has seen and heard after 3 days, and 10% of what is only heard and not seen [16], photographs are of great value as memory aids and evidence. A notable usage of early photography as evidence is the documentation of the crime scenes and victims of Jack the Ripper in 1888 [17]. But this is far from the earliest crime related photography. The use of photographs for mug shots and finger prints goes as far back as 1841 [18], only two years after the first publically announced photographic process by French inventor Louis Daguerre [19]. Photographic enhancements had their day in court not long after the first photographic evidence, with records showing that image enlargements were used in trials since as far back as 1860 [16]. Since these initial incursions of photography and image enhancement into the courtroom, their use has been indispensible in matters of law.

Admissibility

In the United States, the admissibility of evidence in court is guided by the Federal Rules of Evidence (FRE) and legal precedent. The FRE sets the rules for admissibility of evidence in federal courts while state courts have their own rules which are typically modeled after the FRE. Under the FRE, anyone who wishes to put forth evidence for the consideration of a court must first establish its relevance and authenticity. The FRE states that all relevant evidence is admissible, with the definition of relevant evidence being "evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence". Relevant evidence can be excluded per Rule 403 of the FRE on the grounds of it being prejudiced or unfair, misleading, confusing, or wasteful. [20]

The use of photographs and photographic duplicates, including enhancements, is permissible under the FRE as explained in Article X, Rules 1001 through 1004. Rule 1001 takes an open ended definition of photographs that can include items such as X-Rays, video tapes, and motion pictures. Duplicates, including miniatures, enlargements or re-recordings of images, are permissible to the extent that they can be determined to be generated from an authentic original, as defined in Rule 901. Duplicates are not acceptable in lieu of an original unless use of an original is deemed unfair. This means that an enhanced image will most definitely be required to be presented in tandem with an original in court. In situations when and the validity of an enhancement is questioned, it is not unheard of for an expert to be asked to conduct an image enhancement on an original copy live before a jury [21].

Expert testimony

Qualified forensic experts are often called in to give testimony in court as part of their role as teachers in the presentation stage of a forensic investigation. In the adversarial system of law as practiced in the US, each party can secure an expert. However, the expert's role is an impartial one that is not intended to further the interests of either party. The expert's only loyalties should be to science. The findings of an expert are used to assist the trier of fact (usually the jury) in making a well-informed interpretation of the available evidence. To be deemed qualified, experts must be demonstrated to have sufficient knowledge, skill, experience, training, or education necessary to testify on the matter at question. There is no specific amount of training or education that automatically qualifies one as an expert. Matters of qualifications are dependent on a case by case basis depending on the needs of the courts. Moreover, the number of times that an expert has served as a witness, whether great or small, does not qualify or disqualify an expert from serving in future trials. For example, in Bogosian v. Mercedes-Benz of N.Am., Inc., 104 F.3d 472, 477 (1st Cir.1997), the court rejected an opinion of a witness who had testified as an expert 126 times because his expertise did not coincide with the disputed issue [22]. Meanwhile, the court of United States v. Locascio, 6 F.3d 924, 937 (2nd Cir. 1993), concluded that "…even the most qualified expert must have his first day in court" [23].

Unlike a lay witness, the opinion of an expert witness is admissible in court so long as "(1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." [20]. The role of an expert witness and admissibility of expert opinion is outlined in Article VII of the FRE, which in turn was interpreted by The US Supreme Court decision in the well-known trial of Daubert v. Merrel Dow Pharmaceuticals, 509 U.S. 579 in 1993. The outcome of this trial set a precedent for the admissibility of scientific evidence that is followed in all US federal courts and modeled after in the rules of many state courts [26]. According to the Daubert standard, the judge serves the role as a "gatekeeper" by assuring that scientific knowledge presented by a witness is relevant and reliable. In this determination of relevance and reliability of a scientific finding, the following non-exclusive checklist can be considered:

1) Whether it is testable, falsifiable, and if it has in fact been tested

2) Whether it has undergone peer review

3) The error rate if known

4) The existence of maintenance standards and controls

5) The degree to which it has been generally accepted by the scientific community [24]

The Frye test, based on the precedent of Frye v. United States 293 F. 1013, was the standard against which scientific evidence was weighed in US Courts before it was superseded by Daubert at the federal level. The Frye standard advocates "general acceptance" in the scientific community as the standard that scientific theories and opinions must meet in order to be admissible in court. The case for general acceptance was delineated by the court in the now famous excerpt from the Frye v. US legal proceedings:

"Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while the courts will go a long way in admitting experimental testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs." [25]

At the state level, Frye modeled standards continue to be employed in place of the Daubert standard in under half of US States [26]. Nevertheless, concerns that Frye is not open enough to novel uses of science, even if conducted with sound scientific methodology, have caused support of it to waver in recent years. Within the last two years of this writing, both Florida and Kansas have dropped the Frye standard in favor of the Daubert standard [27][28].

Challenges

Digital images and enhancements can face many challenges in court due to real and perceived issues surrounding digital images. Many of these challenges arise from misconceptions of the differences between digital photography and film. Some believe that film cannot be altered and therefore is of better service to the court. This is a myth. The alteration of film goes all the way back to 1840 [13]. Other concerns are that image enhancements will change the content of an image, with the possibility of say, changing a fingerprint from one to another. In truth, if forensic best practices are followed, the probability of this happening has been found to be 1 in 10 to the 80th power according to a study conducted at Purdue University [29]. Another common challenge arises because of perceived problems with the reproducibility of image enhancement. It's true that forensic image enhancement is a heuristic process. For any given image being enhanced by multiple experts, each one will undoubtedly use different sets of enhancements to different extents based on his or her training and experience. Though reproducibility is a necessity of the scientific method, reproducibility is judged by the ability to obtain visually comparable results, not identical ones [30].

The Scientific Working Group on Image Technology (SWGIT) details a large number of myths and facts about the admissibility of digital images and enhancements in courts, like the ones seen above, in their document "Digital Imaging Technology Issues for the Courts" [30], but perhaps it is best to look to legal precedents to get an accurate picture of digital images on trial.

In the Georgia Supreme Court case of Almond V. State, Almond raised doubts about the use of images as evidence on the grounds that they were digital, not film. The court found that the evidence was properly authenticated and admissible. The court went on to say, "We are aware of no authority, and appellant cites none, for the proposition that the procedure for admitting pictures should be any different when they were taken with a digital camera. Therefore, the trial court did not abuse its discretion by allowing the photos into evidence here."[31]

During Washington Court of Appeals case of State v. Hayden, when the challenge of an enhancement of a latent fingerprint though software arose, the court asserted "Because there does not appear to be a significant dispute among qualified experts as to the validity of enhanced digital imaging performed by qualified experts using appropriate software, we conclude that the process is generally accepted in the relevant scientific community." [31]

In Nooner v State 907 S.W.2d 677(Ark. 1995); Nooner v State 322 Ark. 87 (1995), the use of a brightening and digital mosaicking filters (as seen in Fig. 1.1) were permitted after the court found that no features were added or subtracted to the evidence images. [32]

The trial of Rodd v Raritan Radiologic Associates, P.A., et al. 860 A.2d 1003 (2004) is one in which the use of an enhancement was not permitted [32]. The court had originally permitted enlarged copies of medical films. Later it was found that the magnification applied to the films distorted the imaged content rather than clarified it . The experts on both parties agreed that they had difficulty interpreting the images as they were not used to making observations on images magnified to such an extent. In addition, content that appeared sharp in the original was blurred in the enlargement [33]. This resulted in a retrial [32].

One of the most important cases to take note of is that of State V Swinton, 847 A.2d 921 (Levy-Sachs). During the trial both the State and Swinton brought forth experts, each of which presented testimony with the use of computer created image overlays. While the State expert had no trouble explaining how the overlays he used were generated, Swinton's expert had not created the images he presented and was found to not have the computer expertise to assess the reliability of the methods in which the images were created. The court found that experts testifying should be the ones that create the enhancements [18]. This has set a guideline for ensuring that experts have the sufficient knowledge to endure cross examination with regards to the software tools they employ. [34]

From the above, we can see that digital images and digital image enhancements have a place in the court so long as their relevance and reliability can be ascertained. There is no automatically allowable or disallowable enhancement technique. Ultimately, any enhancement must survive a Daubert or Frye test in order to satisfy the admissibility requirements of the courts.

Quality assurance

As explored in the section above, the extent that forensic science can serve justice is contingent on how well it can be demonstrated that the scientific methods used are reliable and the conclusions derived from them are valid. Therefore, forensic scientists have an obligation to perform to a high degree of excellence and forensic laboratories are equally obligated to have a quality management system in place that can ensure and demonstrate that this is indeed so [35]. In media forensics, a quality management system should document permissions, obligations, and prohibitions [35] for laboratories and examiners encompassing four major categories: The integrity of digital evidence, the validation of methods, the verification of tools, and the competence of examiners.

Integrity of evidence, as briefly mentioned earlier, refers to the assurance that evidence is free of damage or contamination through acquisition and throughout processing. For digital multimedia, this includes creating forensic copies of evidence through write blockers to prevent the accidental alteration of data, verifying the copy through a hash function which should yield identical values for the original and copy of a digital file if they are indeed exact duplicates of each other, ensuring the safe storage of evidence, storing a permanent copy of the original data, and maintaining a well-documented chain of custody [36]. For image enhancement, best practices dictate that enhancement should only be performed on a forensic copy [36] and that the enhanced image should not be presented as an original. Rather, it should be made clear that an enhanced image has been processed and detailed documentation should reflect what processing has been done to it.

The validation of methods and tools is conducted by subjecting tools to scientific testing. Valid methods and tools must be demonstrated to be suitable for their intended purpose, based on accurate knowledge, and be able to produce repeatable and reproducible results. Repeatability relates to the ability of a method to obtain same or similar results in the same testing environment. Reproducibility describes the ability of a method to obtain same or similar results in a different testing environment, usually by a different scientific team. After successful validation, periodic verification is expected of all labs to ensure that their tools and methods continue to produce expected results.

The competence of examiners is the foundation that all quality forensics relies upon and the most crucial component for a quality management system. Being that quality examiners are more likely to produce reliable results, examiners must receive adequate training in their assigned duties, undergo competency testing, and undergo periodic proficiency testing. According to the Scientific Working Group on Digital Evidence (SWGDE), experts should at minimum be required to undergo 40 hours of discipline specific training annually, and be well versed in the applicable laws relevant to their roles [37]. In addition, examiners must abide by ethical codes of conduct as enforced by the legal system and by most professional forensic associations such as the American Academy of Forensic Sciences (AAFS) the International Association for Identification (IAI) and the European Network of Forensic Science Institutes (ENFSI). Commonly endorsed codes of conduct dictate that forensic examiners must at a minimum testify honestly, not overstate or understate conclusions, be involved in the scientific community, maintain and update technical skills, observe academic honesty, avoid conflicts of interest, understand limitations of science, and correct errors when found [38].

The Quality Assurance System should be written in a Quality Assurance Manual (QAM) to ensure that whatever is written in the manual is done, and whatever is done is written [35]. The QAM makes the QAS easy to review and revise as needed. Part of the Quality Assurance Manual should list Standard Operating Procedures (SOPs). SOPs are lab specific documents which instruct how to do specific tasks. They are a necessary part of any quality management system as they aid in ensuring that examiners have a protocol for various lab functions such as the handling of evidence, calibration of instruments, and proficiency testing standards. Anything that is carried out that is not documented as a lab practice in the QAM could be considered a misconduct of the lab or examiner and may not be able to stand up to legal scrutiny.

Laboratory accreditation

Labs that wish to increase confidence in their technical competency may seek out accreditation. Accreditation standards of scientific laboratories in many disciplines including forensics are based on *ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories*, the collaborative work of the International Organization of Standardization (ISO) and the International Electrotechnical Commission (IEC). This standard is recognized internationally and is upheld by accrediting bodies such as American Society of Crime Laboratory Directors-Laboratory Accreditation Board (ASCLD/LAB). [35]

Currently, accreditation is not a mandatory requirement for many forensic labs around the US but this may soon change. The National Academy of Sciences' report "Strengthening Forensic Science in the United States: A Path Forward", has made a strong impression in the forensic community since its debut in 2009 [39]. Among the many recommendations that the NAS made to improve forensics in the US, the standardization and accreditation of labs been one that has divided many opinions in the forensic community. SWGDE's position on the matter is that while accreditation is certainly preferred, it is not always a possibility for small labs due to the financial and personnel burden that the process entails [40]. Though unaccredited labs may choose not to undergo accreditation for many reasons, they still have an ethical and scientific responsibility to assure the validity of their methods and analyses by keeping a quality management system in place, one which preferably conforms to the ISO/IEC standards.

Change in the horizon

As of this writing, the state of forensics in the US is going through a period of change. Since the 1990s, forensic best practices and quality standards have been coordinated and developed by subject matter experts from multinational scientific, federal, and educational institutions working together under the purview of the federally funded Scientific Working Groups (SWGs). This has begun to change with the introduction of the newly developed (as of 2014) Organization of Scientific Area Committees (OSAC). The OSAC have been created by NIST with the aid of the US Department of Justice as a direct response to overcoming the deficiencies in the organizational structure in the SWGs and problems in forensics as a whole as outlined by the 2009 NAS report [42]. In the original SWG framework, there were 21 scientific working groups functioning as separate entities. Under OSAC, there are five area subcommittees and each one is a part of a larger whole overseen by the newly created Forensic Science and Standards Board (FSSB) [42]. OSAC is organized into the following five subcommittees: Biology/DNA, Chemistry/Instrumental Analysis, Physics/Pattern Interpretation, Crime Scene/Death Investigation, and Digital/Multimedia.

The future of forensics will hinge on the developments of NIST and the OSACs but as of yet, the OSAC are still in development and no OSAC subcommittee has published any documents establishing quality standards for forensics and they are long way from doing so. In the meantime, the SWGs are still relevant, with SWGIT and SWGDE being the two most pertinent to image enhancement. Though SWGIT has been defunded, their published works throughout the years serve as "living documents" that continue to be available for public comment and critique. These documents constitute the general consensus of the scientific community and continue to be well regarded. SWGDE is one of two SWGs that has not been defunded (the other being SWGDAM which oversees DNA evidence [41]) and will continue to operate outside of the OSAC system, with future plans to expand into former SWGIT territory. In 2015, forensic video and image committees which constituted SWGIT were created within SWGDE, which is now in the process of republishing SWGIT documents (with revisions as necessary).

SWGIT and SWGDE have separately and jointly published a large number of documents for the dissemination by the media forensics community. It would be a disservice to attempt to summarize the content of all of these documents within this paper and it is highly recommended that their documents, especially those presented in Table 1.1, be sought out and read in their entirety from their original source for a more detailed insight of quality assurance and best practices in media forensics.

SWGDE Documents
SWGDE Proficiency Test Program Guidelines
SWGDE Proficiency Test Guidelines
SWGDE Digital and Multimedia Evidence (Digital Forensics) as a Forensic Science Discipline
SWGDE Min Requirements for Quality Assurance in Processing Digital Multimedia Evidence_v1
SWGDE Position on the NAS Report
SWGDE Best Practices for Photographic Comparison for All Disciplines (Draft)
SWGDE Image Processing Guidelines (Draft)
SWGDE Training Guidelines for Video Analysis, Image Analysis and Photography V1-1 (Draft)
SWGIT Documents:
Section 1 Overview of SWGIT and the Use of Imaging Technology in the Criminal Justice System
Section 5 Guidelines for Image Processing
Section 6 Guidelines and Recommendations for Training in Imaging Technologies in the Criminal Justice System
Section 11 Best Practices for Documenting Image Enhancement
Section 12 Best Practices for Forensic Image Analysis
Section 17 Digital Imaging Technology Issues for the Courts
Section 18 Best Practices for Automated Image Processing
Section 24 Best Practices for the Retrieval of Digital Video
SWGIT/SWGDE Joint Documents
Guidelines and Recommendations for Training in Digital and Multimedia Evidence
Recommended Guidelines for Developing Standard Operating Procedures

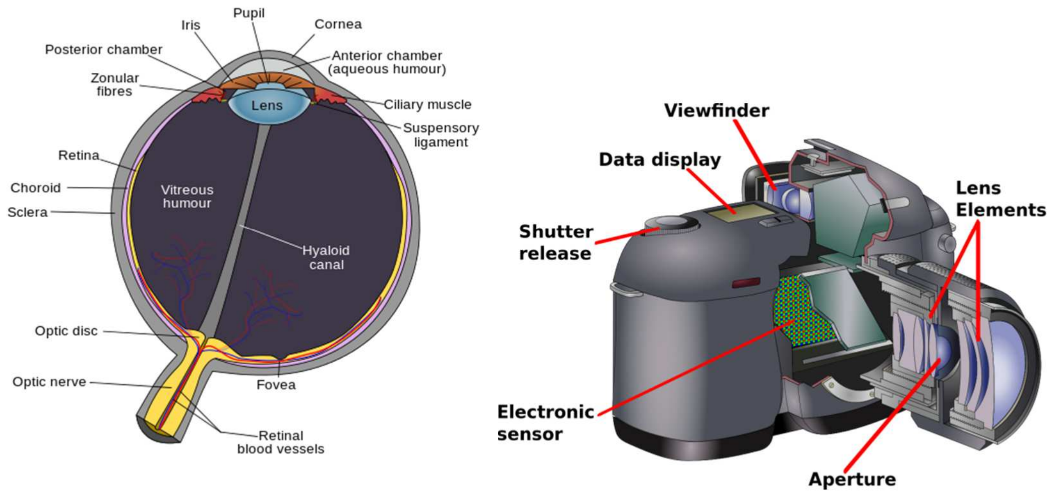
Table 1.1: A collection of SWGIT/SWGDE documents

CHAPTER II

FROM LIGHT TO PIXELS

A typical camera serves the purpose of capturing information for the eventual delivery to the human eye. With this ultimate goal in mind, manufacturers develop cameras with specifications which correspond with the limitations and strength of the human visual system. Naturally, for most uses it would be wasteful to build cameras which could capture minute differences in color or detail that are imperceptible to humans. Even with cameras sensitive to information that surpasses human abilities, like infrared or night vision, this information is generally translated after capture into something that can be readily interpreted by humans. Despite the inclination toward creating useful images for human consumption, the typical camera can and will capture information that is not readily suitable to the human visual system without additional image processing. This close interaction between properties of cameras, images, and human vision is at the core of any enhancement operation and is deserving of being studied by forensic image enhancement professionals.

Human vision is in many ways analogous to a standard digital camera. Of course this is no accident. Since the beginning of humanity, inventers have taken inspiration from the wonders of the natural world, and therefore it makes sense that camera manufacturers construct cameras based on the *construction* of the human eye. A simplified comparison shows that like most any camera, the eye has a lens, aperture, light sensor, color receptors, and an image processor; these are the lens, iris, retina, rods and cones, and brain respectively (Fig. 2.1). Through in reality the human visual system is much more complex than even the most advanced camera [43], this analogy serves well for exploring the connection between sight and the modern camera.



(a) (b)

Figure 2.1: Diagram of a human eye (a), Diagram of a digital camera (b)

The human visual system

Contrary to conventional wisdom, humans don't see with our eyes; we see with our brains. That is to say that our eyes are merely light sensors that transfer information to our brains, at which point that information is processed into a (hopefully) accurate picture of the world. In fact, it is argued that the optic nerve, and sometimes even the entire eyeball itself, is simply an extension of the brain [18]. Different states of the brain such as those induced by stress, medication, injury, etc. can have very real effects on what one sees. Indeed, most everyone has been so absorbed in though at one time or another to not process the visual information in front of their eyes only to be jostled back in to the world of sight by a loud sound or a bump in the road (Hopefully not both at the same time) with no memory of what passed before their eyes. In other words, sight is both a physical and psychological phenomenon [18].

The connection between brain and sight is no sooner apparent than in the discussion of color and intensity perception. What the human eye registers as visible light is in fact high frequency radiation. This radiation is a sliver of the electromagnetic spectrum with a wavelength of between roughly 400 - 700nm (Fig. 2.2) [43]. Color receptors called Cone cells, found in the back of the eye in part of the *Retina* known as the *Fovea,* are sensitive to a certain wavelengths within the range of visible light which the brain interprets as different colors. There are three types of cones known as S,M, and L, in reference to the short, medium, and long wavelengths that they are most sensitive to. S cones are most sensitive at wavelengths roughly 420-440nm (blue), M cones are most sensitive in the 564-580nm wavelength range(green), and the L cones are most sensitive in the 564-580 nm range (red) [44]. In order to see colors that do not correspond to the wavelengths detected by the cones of the eye, the information gathered by the three cones types must be sent out by the *Optic Nerve* to be combined and processed in the brain [17.] For example, information received from green and red may combine for the perception of the color yellow. Information received from all three cones can also be used to perceive colors with no corresponding wavelength in the visible light spectrum like pink. The connection between the brain, wavelengths, and colors is encapsulated perfectly in the words of authors Blitzer, Stein-Ferguson, and Huang: "Wavelengths do not have 'colors'-humans do." [18]

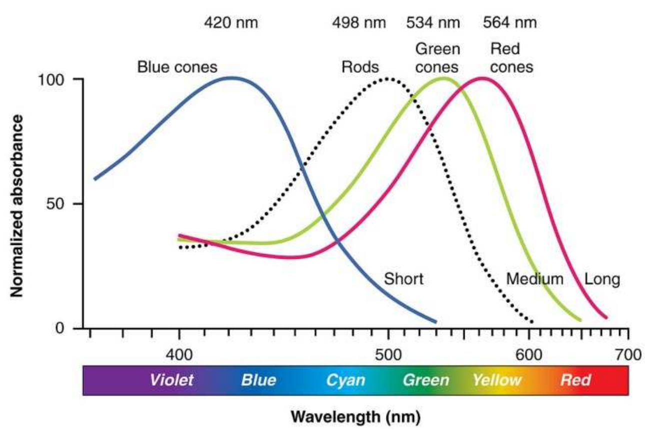


Figure 2.2: Human color response

The retina also holds cells called Rods, which are sensitive to the green/yellow range of visible light and particularly receptive to changes in light intensity and low light. Rods move forward within the eye during low light in order to aid in night vision [18]. While rods make no distinction between colors, they do affect overall perception of color. Human perception of color and brightness are not based on linear measurements but rather on comparisons and expectations of the world [21]. This means that the color perceived by the eye is affected by brightness of the scene observed and vice versa. This effect is the cause of the optical illusion seen in Figure 2.3, in which the exact same color and brightness values appear to be different based on the content of the scene. It has been shown that the human visual system can discern an estimated 26 levels of brightness in a scene at a time [15], with changes in brightness only being noticeable when they are between 2 and 3 percent higher than the value being compared to [21]. Because human color and brightness discernibility is based on an increased percentage of a compared value, human color and intensity perception is understood to be logarithmic rather than linear [21]. Camera sensors on the other hand do capture color and intensity information linearly, with changes in brightness corresponding directly with the amount of visible light.

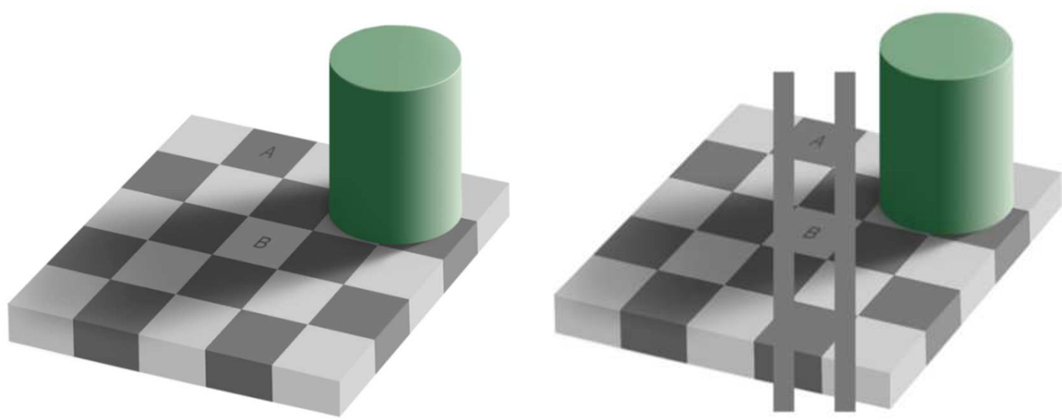


Figure 2.3: Intensity difference illusion

Notice that tiles A and B are in fact the same shade of gray.

All together the rods and cones make up somewhere in the vicinity of 110 – 150 million sensors per eye [21][43]. By comparison, a new consumer grade digital camera in the market as of this writing may contain roughly 12 million sensors and standard 35mm film can display approximately only 5000 distinct points of color [21]. It's clear that in terms of the sheer amount of individual color components that can be captured and displayed, the human eye is the clear winner. The human eye is indeed a marvel of nature but it has its blind spots, figuratively and literally. The blind spot of the eye, caused by a lack of rods and cones on the optic disc[45], is a defect of the eye found in humans as well as other vertebrates, though not present in Cephalopods, like the octopus, whose eyes are the result of a different evolutionary path form our own[46]. The blind spot is not easily noticed as the human brain fills in the missing visual information with an approximation based on what is being observed. The location of the blind spot in humans can be detected however, by the use of a simple test as image as in Figure 2.4. This short discussion of the blind spot is briefly mentioned here to illustrate both the psychological nature of our vision and to highlight its limitations.

RIGHT EYE

LEFT EYE

A blind spot can be identified by closing one eye and focusing the open eye on the text corresponding to the eye that is open. By moving forward or away from this figure, the text at the periphery of the open eye will disappear at approximately 3x the distance between the red and green text boxes.

Figure 2.4: Blind spot detection test

One such limiting trait of the human eye is its limited perception of color variance in relation to the relatively high perception of intensity. This phenomenon is a result of cones making up approximately only 5% of the total light sensors in the eye, with the rest of the light receptors being rods [47]. While the human visual system has great response in regards to green light [21], color perception is not as strong for red light and extremely limited in the blue range as only about 2% of the cones in the eye are S cones [47]. Human visual acuity, the ability to resolve fine details by sight, has its own set of limitations. If we were to imagine a set of parallel lines gradually shrinking into a vanishing point but never crossing, there would be a point where the human visual system could no longer discern the lines as individual lines but rather see them as one solid line. A human with 20/20 vision can resolve no more than 5-6 lines per millimeter at a comfortable viewing distance [43]. In the next sections of this chapter we will see how such properties of the human eye are exploited for image creation and enhancement.

Creating a digital image

It is generally presumed that photographed images begin at the camera lens, but let us stop and consider the purpose of a lens. A precursor of the first photographic cameras known as the camera obscura (Latin for “dark room”) worked without a lens, based on the pinhole camera phenomenon. This phenomenon, whose first record of discovery goes back to 5th century BC China [48], refers to the projected image created as a result of light passing from a small hole into a suitably dark area. The astute may ask, "Why then is a lens necessary if an image forms without one?" The answer to that question is sharpness and intensity. The opening, or aperture, which light passes through in this method is not a lens in the traditional sense, in that rather than focusing light, it merely funnels it. The aperture in a camera obscura creates a barrier that only allows light to enter from a narrow angled trajectory. Making the aperture smaller makes the image appear sharper but because the entrance for light is severely restricted, the resulting image created is dim. Making the aperture larger can make the image brighter, but because the angle of the trajectory of light is also larger, light that enters the hole does so at many more angles. As a result, the light corresponding to one point of an object in 3D space reaches multiple points on the projected surface, thus creating what is known as a blur. If the hole is even larger, the blur becomes so large that the multiple colliding rays of light simply combine into a solid colored light due to the additive properties of light.

In an additive color model, colors combine to create an increase in overall luminance (light intensity), as opposed to *subtractive* color, like that of printer ink, in which colors become darker when combined. By introducing a well-focused lens to a camera, light incoming from multiple directions can be focused in such a way that multiple beams of light from one point of a 3D object strikes one point on the projected image. Thus, the resulting image will have both sharpness and intensity. Blur, sharpness, and light intensity are governed by these basic principles in both modern lensed cameras and the human eye.

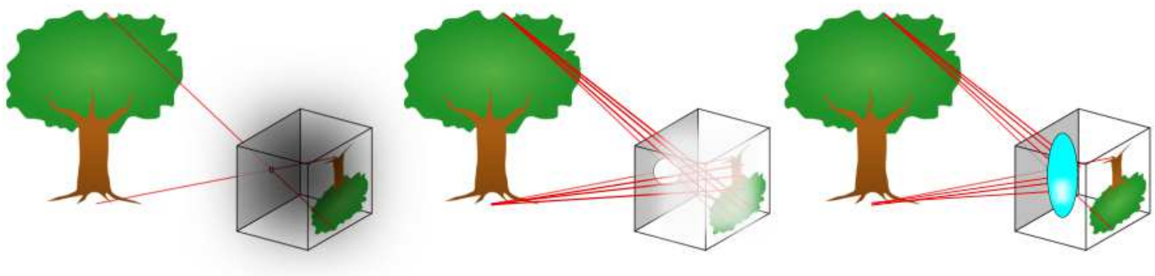


Figure 2.5: Comparison of camera aperture size and lens
Left: Small aperture, Sharp/Dim; Center: Large aperture, Blurred/Bright; Right: Lens, Sharp/Bright.

In order for a photograph to form, the light that passes through an aperture must meet a photosensitive surface for the projected image to imprint on. Where the eye has a retina and an analog camera has a photosensitive film, light flowing into a digital camera is captured by an electronic sensor. The two most common sensors in digital cameras are the Charged Coupled Device (CCD) and Complementary Metal Oxide Semiconductor (CMOS) [43]. Theses sensors have several differences in their function and resulting image quality, but for now, their function can be thought as equivalent. These chips capture particles of electromagnetic radiation known as photons through a matrix of photo sensitive diodes called pixels [21]. The intensity of light reaching these pixels is converted to electrons whose electronic voltage is measured and converted into binary values that can be interpreted as a digital image. The number of pixels wide by the number of pixels high of this matrix is known as the resolution of the camera. A camera's resolution relates to the maximum pixel resolution of the final image, though larger or smaller image resolutions are possible depending on the signal processing applied to an image after capture. Because electronic sensors use a finite number of pixels to capture an infinite number of details in a scene, image data must be truncated to conform to the limits of the camera sensor. This is known as Sampling Quantization. All digital images will have a degree of information loss directly proportional to the size of the sensor matrix, with larger sensors generally being able to generate better representations of a photographed scene. (Fig. 2.6).

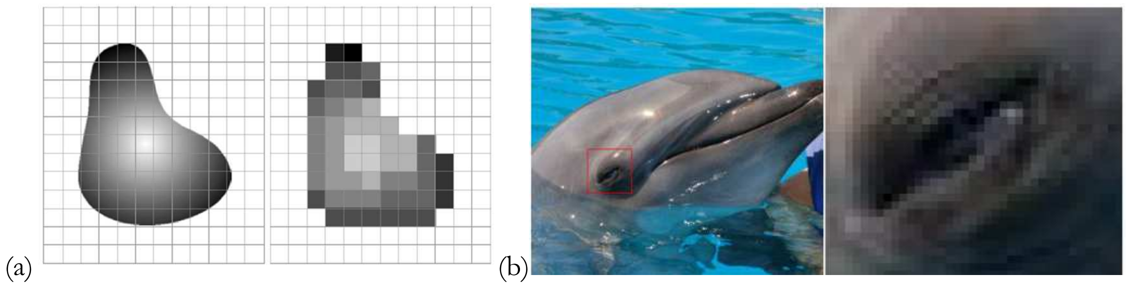
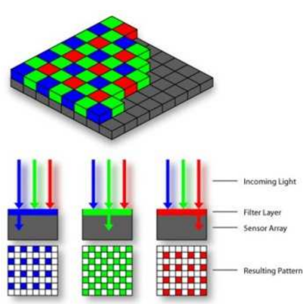


Figure 2.6: Sampling quantization cause (a) [49] and effect (b)
(a)-Left: Original scene as projected on an imaging sensor. A-Right: The result of the left image being quantized by an imaging sensor
(b): Crop showing effect of Sampling Quantization.

While in the above paragraph the words photons and light are used seemingly interchangeably, there is one major distinction. All light is made of photons but not all photons are light. What is meant by light is the part of the electromagnetic spectrum between about 400 - 700 nm in wavelength that is visible to the human visual system. This distinction is significant for the capture of color images. CCD and CMOS censors are naturally sensitive to photons (not only visible light) traveling in the Infrared range of the Electromagnetic Spectrum (a range not ordinarily visible by humans). The imaging sensor consequently cannot see in color unless an Infrared-Cut filter (IRC) is used to see visible light with the same spectral sensitivity as humans [43].

Another factor for creating a color image is the Color Filter Array. Pixels in an electronic sensor recognize intensity of light but not color. The Color Filter Array (CFA) is a mosaic of red, green, and blue color filters attached to the electronic sensor of a camera which attributes the light captured at each pixel element of a sensor into one of the three wavelengths corresponding to a primary color, reminiscent of the function of the cones in the eye. Color filter arrays come in many varieties but the most commonly used CFA is known as the Bayer filter, after its creator Bryce Bayer of Eastman Kodak [50] (Fig. 2.7(a)). Bayer's filter is designed with the understanding that human vision is most sensitive to green [21]. 50 percent of pixels in this filter are green while the other 50 percent are divided between blue and red. When light is captured with the use of a CFA, each pixel only records information for one color. For the final color to appear natural in the output image, the other two colors that are missing in each pixel must be extrapolated from the surrounding pixels through a process called demosaicing (Fig. 2.7 b). The simplest method of demosaicing is to generate the missing color values of a given pixel by averaging the color values of its neighboring pixels. This demosaicing method generally produces a good approximation of the colors of a captured scene since adjacent areas of a scene will typically share the same or similar color values. This reasoning breaks down at edges of objects however, where color information is not closely related at the edge boundary. This causes demosaicing extrapolation error which can create extraneous spots of color around edges [50] (Fig 2.8). Because of this, many camera manufacturers use more advanced demosaicing algorithms in an attempt to improve the quality of their images. High end cameras used for television broadcast get around the problems of demosaicing by having three full sensors for each color inside of one camera. This technology has not reached the general public due to the high costs and bulk of these cameras.

(a)


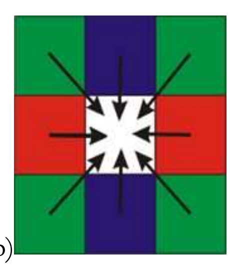
(b)


Figure 2.7: (a) Bayer filter, (b) Bayer demosaicing

Figure 2.8: Bayer demosaicing "zipper" artifacts on edges

Demosaicing is just one part of the penultimate step of the digital image capture pipeline known as Digital Signal Processing (DSP). DSP includes any processing of the signal after it has been captured by the camera sensor. DSP begins with an Analog to Digital Conversion (A/D) which translates the voltage information captured by the imaging sensor into digital values. At this step, quantization takes place to reduce the infinite levels of color and intensity values that strike the imaging sensor into a finite set of intensity values that camera can use. Once these numbers are generated, the camera's internal image processor may manipulate the data in many ways including, denoising, making color adjustments, demosaicing, and more or none of the above before eventually converting the image data to a file type suitable for viewing or storage. [43]

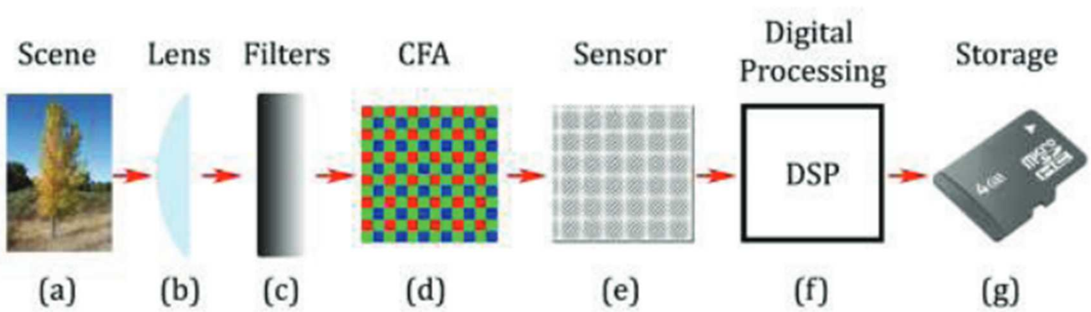
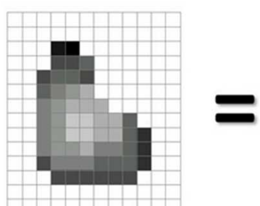


Figure 2.9: The digital image creation pipeline [13]

Properties of digital images

The files stored by a digital camera contain the digitized tonal information captured by the imaging sensor. This information is stored as one or more matrices of numbers whose coordinates in the grid correspond directly to the x/y pixel coordinates of the digital image. A pixel in this context now refers to the smallest unit of color available in an image and not a camera sensor element. The numerical values in these matrices are interpreted to give each pixel its tonal value in an image. Grayscale images contain one matrix for tonal information while color images in the RGB color space will contain three: one for the red, green, and blue, color layer. The values in these matrices describe the intensity of the tonal value captured by the camera sensor. Similar to the processing of rod and cone information by the brain, by combining the information from all three color layers, known as color channels, an image with full color representation can be formed.



255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	20	0	255	255	255	255	255	255	255
255	255	255	75	75	75	255	255	255	255	255	255
255	255	75	95	95	75	255	255	255	255	255	255
255	255	96	127	145	175	255	255	255	255	255	255
255	255	127	145	175	175	175	255	255	255	255	255
255	255	127	145	200	200	175	175	95	255	255	255
255	255	127	145	200	200	175	175	95	47	255	255
255	255	127	145	145	175	127	127	95	47	255	255
255	255	74	127	127	127	95	95	95	47	255	255
255	255	255	74	74	74	74	74	74	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255

Figure 2.10: Array of tonal values for a grayscale image [51]

The range of discreet tonal values that can be recorded and displayed in a digital image is determined by the image's bit depth. A bit is the smallest unit used in a computer's binary numerical system and can have a value of either 0 or 1. The term Bit Depth expresses how many groups of bits can be used to define a certain value, in this case, the individual tonal values per pixel in an image. This relationship between bit depth and tonal range can be described by the expression $L=2^K$, where L is the maximum number of tonal values available and K is the number of bits available for storing the range of tonal values in an image [52]. The most common bit depth for digital images is 8 bits for a total of 256 total possible tonal values per pixel for each color channel. These values are grouped in a range from 0 – 255 where 0 describes no intensity (the darkest black) and 255 describing the highest possible intensity that can be recorded in an image (the lightest white). Because 8 bits can be used to define each of the three color channels, 8 bit images are occasionally (and confusingly) referred to as 24 bit images, not to be confused with actual 24 bit *per channel* images which can contain 2^{16}, or over 16 million, gradients of color as opposed to the standard 256 (Table 2.1). While larger and smaller bit depths are possible, 8 bit images are most common considering that by combining the 3 color channels of 256 values, 8 bit images can contain up to 16,777,216 unique color values, considerably more than the theorized 10,000,000 colors that the human eye is capable of seeing [53]. Because the bit depth defines the number of colors available in an image, bit depth is also commonly referred to as the color depth of an image. With this idea of depth in use, a digital image can be thought of as being a 3D object with x,y coordinates describing the spatial location of pixels and the z coordinate describing the color depth (Fig. 2.11).

Bits Per Pixel	Number of Colors Available	Common Name(s)
1	2	Monochrome
2	4	CGA
4	16	EGA
8	256	VGA
16	65536	XGA, High Color
24	16777216	SVGA, True Color
32	16777216 + Transparency	
48	281 Trillion	

Table 2.1: Bit depth/color range comparison [54]

Figure2.11: Image as a 3D object.
The left image is displayed as a 3D object on the right, with intensity is being interpreted as height.

Digital image files come in a variety of file formats. File formats are standards for organizing, managing, encoding, and decoding information in a computer. The standardization of these formats is what allows files to be distributed with the confidence of knowing that the receiver will be able to open them. For images, they indicate where to find various items of image information and how to interpret pixel data. A format must specify an image's bit depth, resolution, compression scheme, encoding algorithm, and more in order for the image to be viewable. Image formats can be thought of as containers where the outside of the container containing non image data regarding image properties and the inside of the container containing the actual visual image data. The non-image data can include information such as the format header, image resolution, color depth, camera settings used to create the image, encoding/decoding information required so that the image can be properly displayed by a computer, gps coordinates relating of where the picture was taken, and more. This data can be of great value for forensic purposes. While it is out of the scope of this paper to go into detailed usage of this data non image data, for the topic of image enhancement it is necessary to mention that any processing of an image may change or delete this data irrecoverably. Therefore best practices require all enhancements to be performed on forensic copies of digital evidence rather than an original file so as to not disturb the provenance of the original digital evidence.

All formats have their own unique specifications. The specifications of each format type indicate how data is stored, accessed, and compressed while also setting constraints on image properties such as the maximum allowable color depth and color channels allowed. There is no single "best" format as all formats have advantages over others for specific use cases.

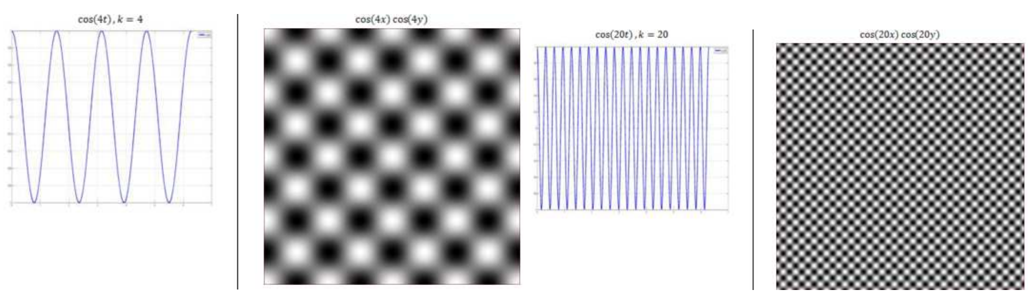
A major consideration for choosing an image format is its capabilities for image compression. Some image formats allow image data to be compressed in order to take up less space in a computer's storage device. Depending on the image format, compression may be lossless or lossy. Lossless compression compresses image data in such a way that it can be uncompressed at a later time without degrading the original image information. If we were to compare the process of storing an image file to the process of packing a backpack, lossless compression would be analogous to carefully packing all your textbooks and school supplies carefully so that they take up the least amount of space. Lossy compression removes information from an image in order to make it smaller. To extend the backpack analogy, this would be like tossing out one of the textbooks from the bag into the dumpster. Lossy compression schemes have an advantage in file size reduction over lossless compression, often being able to achieve compression ratios of 50:1, 100:1, or more, depending on the compression encoding used [21]! Lossy compression formats are generally not used to re-encode images used for forensic analysis, referred to as Type II images by SWGIT, but they may be suitable for Type I images, which are used for documentation [1]. However SWGIT accepts that analysis of lossy compressed image is permissible if a trained examiner can determine that the compression does not remove any pertinent information [1].

JPEG, TIFF, and RAW are the file formats most commonly used by digital cameras today so it is worth taking time to discuss their differences as they relate to image enhancement.

In common usage, JPEG describes a type of file container and lossy image compression developed by the Joint Photographic Experts Group in 1992. In the spirit of accuracy, a distinction should be made between the container and compression method. JPEG is actually a type of compression and JFIF (JPEG File Interchange Format) is the container in which JPEG compressed image data most commonly resides. This distinction is important because JPEG encoded data can reside in other container types but JFIF containers can only carry JPEG encoded data. The JPEG encoding standard supports both 8 bit and 12 bit color per channel [56], though 8 bit is the most commonly used. 12 bit color JPEG has not been generally adopted by the general public and is relegated to specific uses by medical equipment [56]. JPEG compression is achieved by intelligently eliminating information from an image that would not be readily visible to the human eye.

While a deep analysis of the mathematical intricacies of this compression is not necessary for the purposes of this paper, a basic explanation of the underlying processes of JPEG compression is useful for understanding image artifacts resulting from strong compression.

JPEG compression is reliant on altering the frequency information of an image. An image is said to have higher spatial frequency as small details within the image increase (Fig. 2.12).



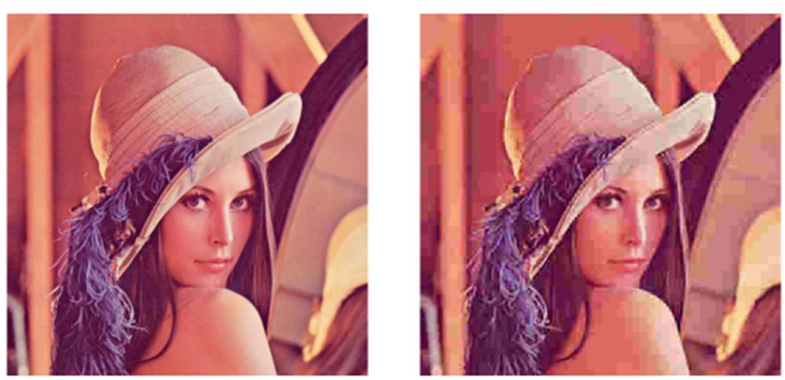
(a) (b)

Figure 2.12: Relationship of image spatial and frequency domain [55]
(a)Long wavelength and corresponding low spatial frequency image
(b)Short wavelength and corresponding high spatial frequency image

JPEG compression begins by converting red, blue, and green (RGB) color information in an image into brightness and blue/red chrominance channels (YCbCr). Because the human visual system can distinguish changes in brightness better than changes in color, this conversion allows the JPEG compression to reduce the amount of color information needed to display an image while leaving the brightness information intact, resulting in a loss of information without a perceivable loss of quality.

To do this, the image is first partitioned into 8x8 pixel groups. The 8x8 groups are then converted into frequency information with the use of a mathematical operation known as a Discrete Cosine Transform (DCT). Once in the frequency domain, the amplitude of the frequency information that makes up each 8x8 area of an image is quantized. Because the human eye is more responsive to low frequency information, the amplitude of the high frequencies corresponding to fine details and color differences imperceptible to the human eye can be quantized using less precision than the low frequencies [18]. After the high frequency information is curtailed, the image is re-converted to the spatial (pixel) domain and stored in a JFIF container. The severity of the compression can often be defined by a user, with stronger compression resulting in higher JPEG artifacts.

Common JPEG compression artifacts include blocking (Fig 2.13), loss of detail, and color distortion. Blocking is the appearance of square areas of color in a JPEG image. These areas are a result of the 8x8 pixel blocks used during JPEG compression. Because each 8x8 block of pixels is processed individually, JPEG compression will create visible edges between bordering 8x8 blocks. Loss of color information and detail is the result of the frequency quantization that corresponds to these details. Smooth areas with less small details show less distortion artifacts after compression. Furthermore, how details are altered is dependent of whether the pixels that make up a detail lie in the center or boundaries of the 8x8 block [21].



(a) (b)

Figure 2.13: JPEG compression artifacts
(a) Uncompressed image, (b) Heavy JPEG Compression

The Tagged Image File Format, or TIFF, is an image file format container than allows for various compression methods. TIFF images can be uncompressed, losslessly compressed or even compressed with lossy JPEG compression. The TIFF specification allows for a higher color depth than JPEG with up to 32 bits per pixel [56]. It can also support a transparency channel, known as an alpha channel, in addition to the standard three color channels found in other formats.

The term RAW does not refer to a single image format but rather a large number of distinct formats used to store the not-yet demosaiced pixel values recorded by a camera sensor. Because the values stored in a RAW file have not been processed into a visible image, it would be slightly inaccurate to call it a true image format. Rather, it is more accurately defined as a data format. RAW files are generally found in high-end and "prosumer" level cameras where users may be interested in having more control regarding the processing of their image files. RAW files offer a great amount in flexibility for image processing and are often equated to negatives in film cameras.

While RAW is often the preferred working format for photographers and enthusiasts, it poses certain challenges for forensic applications. Some may initially see RAW as the perfect format for forensics since the data contained within has the least amount of error possible as none has been introduced through processing. However this line of reasoning breaks down in practice. Even if the data is unprocessed, it serves little use until it does become processed. In addition, RAW formats are not standardized or well documented. Since RAW formats are often proprietary to a specific camera manufacturer, there may be a need to acquire specialized software from the manufacturer in order to read the file. Many camera manufacturers only support specific RAW formats for a short span of 3-4 years before becoming obsolete, making it impossible to process RAW files years down the line. This wouldn't do for forensic applications since evidence must often be archived for 10 years or more. While some programs have the ability of opening RAW files from a variety of manufacturers, it is possible that more processing errors can be introduced if the file is opened with third party software than if it had been opened with the proprietary manufacturer's software. If one is not able to see the imaged content of a RAW image, it serves as a poor piece of evidence in most cases. A format such as TIFF would serve better for archiving and evidence purposes but the file size could be potentially much higher than the RAW as a TIFF would have 3 color channels per pixel as opposed to the one color channel per pixel of the RAW. [18]

Properties of digital videos

Digital video formats follow the same basic properties of color depth, resolution, and formatting as discussed above with digital image files. After all videos are nothing more than a collections of images played back in sequence. The difference is that video container formats have the additional abilities of storing multiple images within a single file container as well as audio.

Because video must display a certain rate of images per second (usually 30) in order to give the illusion of motion, the storage space and bandwidth requirements of video are many times higher than those of still images. In response to this dilemma, a large number of video compression methods are available. Videos formats can make use of still image compression techniques by treating every frame as an independent image in what is known as "Intraframe coding". For example, every frame of video encoded in the Motion JPEG format (MJPEG) is compressed using standard JPEG compression. More efficient compression schemes exist that take account of the redundant temporal data between frames. For example, imagine a video of a car race where a car is speeding through a road. If the car is moving across the road but the road is still within the video frame, the compression algorithm will treat the background road data as redundant and only store the data needed recreate movement of the car. This is known as Interframe coding. Interframe coding uses a Group of Pictures structure (GOP) where the following frame types are possible:

I frames: Intracoded frames containing the full scene information.

P frames: Predictive frames generated based on the previous I frame or P frame.

B frames: Bidirectional frames which are generated based on the previous and next I or P frames (Fig. 2.14).

This method of using redundant and temporal frame data is at the framework of the commonly used H.264/MPEG-4 AVC video coding standard [57].

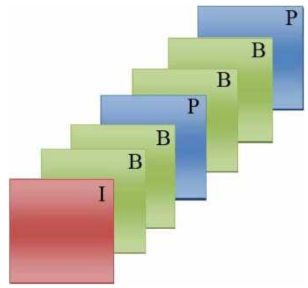


Figure 2.14: Group of picture video frame structure

Image quality and enhancement potential

The properties and limits of the eye, the camera, and the image file are all factors influencing the *quality* of a digital image. In this usage, quality describes both the actual visibility of the image content and the degree to which it may be enhanced. This section unifies the information covered in the previous sections of this chapter in order to examine what factors influence image visibility and enhancement potential.

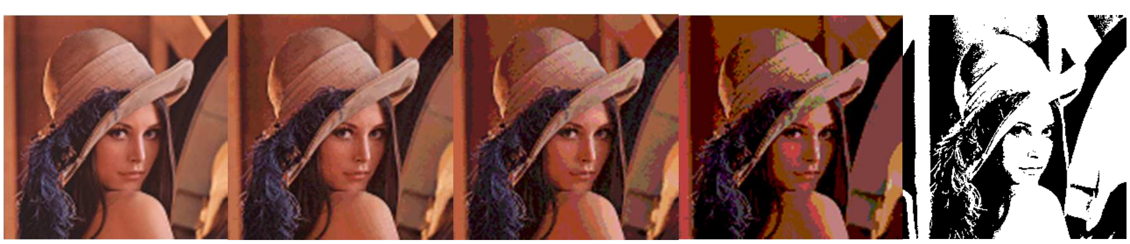
The limits of image enhancement are largely determined by the amount of information loss accrued through each step of the image creation pipeline (Table 2.2). The sampling quantization caused by the image sensor's limited resolution as seen in section 2.2 is the first culprit of information loss. In science fiction, endless detail can be achieved by zooming into an image. In reality, the actual limit of the size of detail that can ever be discovered in an image is set early on by spatial dimensions of the imaging sensor. Spatial quantization determines the final resolution of an image. Enlarging a single image cannot bring out details that surpass the original pixel limit of the sensor. Instead it will only make the appearance of individual pixels larger. Though enlargement does not add detail, it could aid someone to resolve details when limitations in their eye's visual acuity prevented them to do so.

Possible causes of image quality loss in the image creation pipeline	
Scene	Camera shake (blur)
	Object motion (blur)
	Low light
Lens	Out of focus
	Lens distortion
CFA	Color layer information missing at each pixel
Sensor	Noise
	Spatial quantization
	Dead Pixels
DSP	Quantization error
	Compression artifacts
	Bayer extrapolation error

Table 2.2: Image quality loss in the image creation pipeline

Even if a camera sensor is capable of capturing a high spatial resolution, this does not guarantee a high optical resolution. Optical resolution describes the maximum amount of detail that can be resolved in an image based on everything used to create, process, and view the image. This could include the camera lens, the camera's internal resolution, the screen used to display the image, and even the eyes of the observer. If the camera has a great internal resolution but the lens is not well focused, the optical resolution will suffer. Likewise, if both the camera sensor is of superb quality and the lens is perfectly focused, the final image result may still show loss of optical detail due to image compression and processing artifacts like the noise introduced through demosaicing and attenuation of detail in strong JPEG compression. Certainly, many of these optical resolution limits can be improved with the use of denoising, sharpening, and deblurring techniques. However, the effectiveness of these techniques is largely dependent on the spatial resolution of the image and the degree of information loss. Depending on the severity of JPEG compression for example, high frequency information in an image will be either attenuated or completely removed. If it's the former, image sharpening could increase the amplitude of the high frequency information and make it more visible. If it's the latter, no amount of image processing can bring the information back.

This concept of information loss extends to the range of tonal intensity values used to display the image. Because the eye perceives intensity logarithmically and cameras record intensity linearly, some tone differences captured by cameras will be too small for the eye to discern. As long as there is at least one level of intensity difference between two recorded pixel values, their differences can be made to stand out and this will be perceived as an image detail (Fig. 2.15 e). However many factors in the image creation pipeline cause distinctions between pixel values to disappear. Due to quantization rounding errors, either through an A/D conversion or image compression, intensity values needed to show details in an image will be rounded to fit the nearest integer in tonal/color range. This quantization error is magnified when images are saved with a low bit rates due to a smaller range of colors that values must be mapped to. Loss of tonal information can also occur when the recorded or processed values of an image far extend the maximum possible limits of the recording medium. This effect is called "clipping" and can be recognized by areas of an image consisting of a solid color made up either the lowest or highest possible tonal range allowable for the image. This occurrence is common for most cameras when attempting to capture scenes with a high dynamic range of tones.



(a) (b) (c) (d) (e)

Figure 2.15: An image from a high (a) to low (e) bit depth.
E shows a binarized image with that can only display one of two tones. Even with only two tones, image features are distinguishable.

Factors that determine image quality in digital video files are the same as those in digital images. Video files can be processed just as images by decoding their frames and treating them as still images. However video files do have some advantages and disadvantages over still images. Though digital videos can theoretically have the same image quality as digital images, in practice the quality of individual frames in digital videos is often lower than that of still images due to high bandwidth and storage capacity needed for high definition videos. However, many enhancement techniques exist that can use information from multiple frames of video which often surpass the enhancements that are possible with a single image.

As we can see, image enhancement is a sort of paradox. For good enhancement results to be possible, the image must be of good quality to begin with. Unfortunately, the quality of digital image evidence is often too poor for image enhancement to be of much help. The types of surveillance systems used by many small businesses often sacrifice image resolution in order to save on storage space. In addition, the small recording devices which are commonplace in covert investigations often produce less than ideal image quality. Though image enhancement in the wild may not always perform as desired, the applicability of the techniques presented in the next chapter will only increase with the advent and adoption of higher quality imaging devices. That being said, even if image enhancement can only reveal a small amount of extra detail, any extra information that can be extracted from a piece of digital evidence, no matter how small, can make a world of difference to a forensic investigation and can be the deciding factor for determining if someone's civil liberties are taken or restored.

CHAPTER III

A SURVEY OF COMMON IMAGE ENHANCEMENTS

The following is a collection of the most commonly found image enhancement techniques. While this list does not exhaust all possible enhancements, the ones presented here are generally well known and accepted within media forensics.

Image enhancements can take place in either the spatial or frequency domain. That is, the pixel data can be modified directly or it can first be transformed so that it can be processed like a frequency signal, as is done during JPEG compression. This is commonly done with a mathematical process called a Fast Fourier Transform which results in low spatial frequency information in an image to become long wavelength information, and high spatial frequency image content to become short wavelength information. The distinction between these two types of image processing methods is not always apparent as all the mathematics involved with either method are usually handled internally by most image processing software and because there are often equivalent methods of accomplishing different enhancements through either the spatial or frequency domain. Nonetheless, it is a useful concept to keep in mind as we explore the underlying processes of the image enhancement techniques listed below.

Resizing

Image resizing refers to the processes of increasing or decreasing the size of an image. Decreasing the size of an image involves removing pixels from an image and degrades the quality of an image. As a result, decreasing the size of an image is not a recommended forensic practice. When two images of disproportionate sizes must be compared, it is would therefore be best to increase the size of the smaller image instead of decreasing the size of the larger. Alternatively, image enlargement is one of the most commonly used image processing techniques applied to forensic image evidence.

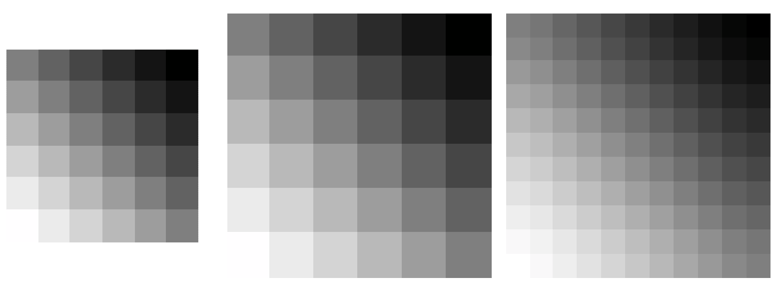
Images are commonly resized in forensics for the purpose of measurement, comparison, analysis, and presentation. For instance, a typical usage of magnification for facial identification may entail increasing the size of a suspect's face in an image so that it may be compared to a reference image of a larger size. The result of the enlargement could then be enlarged again to serve as a visual aid for a courtroom.

Image enlargement generally increases the spatial resolution of an image through processing in the spatial domain. For enlargement operations to increase the spatial resolution of an image, additional pixels have to be generated through interpolation of the existing pixels. The simplest and fastest way to do this is to increase the pixel matrix size of the original image, leaving space for new pixels to be inserted, and filling the blank spaces with the same pixel values of neighboring pixels. This method of interpolation is called Nearest Neighbor interpolation. Though additional pixels are added with this method, it is done in such a way that no new visual information is created and it is best thought of a redundant or duplicative representation of information as opposed to an increase of data. While this method of interpolation may seem ideal for forensic use since no new data is created, the strong pixelation effect caused by effectively increasing the size of each pixel in an image may in itself be perceived as a misrepresentation of image features if the magnification is very severe.

Other methods such as Bilinear and Bicubic interpolation reduce the blocky appearance of image enlargements by generating new pixels through more advanced mathematical means. The bilinear method begins like the nearest neighbor method by increasing the matrix size of an image but rather than replacing empty pixels with the exact same value of a neighbor, a new pixel color value is estimated by averaging together the values of neighboring pixels. Bicubic interpolation is a more complicated and computationally demanding method than the bilinear interpolation but it is more sensitive to image details [18]. Instead of generating new pixel values from four surrounding pixels like in bilinear interpolation, bicubic interpolation uses data from 16 pixels, with pixels closest to the unknown pixel having a greater impact on the calculation. This results in an enlarged image that is a high quality, sharp, and free of most enlargement artifacts [58]. The drawback of bicubic interpolation is that it is more time consuming than either nearest neighbor or bilinear methods. While this may not be an issue for a single image and a modern computer, it may not be a time effective method to enlarge a high quality video.

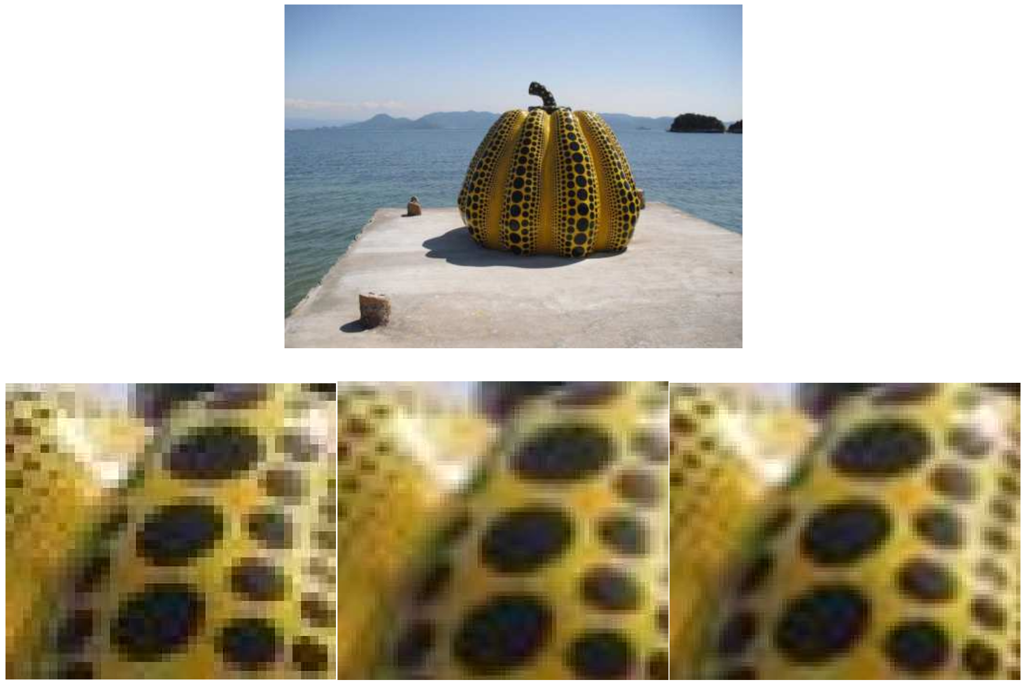
A common misconception with image enlargement is that enlarging an image will increase detail. Though enlargement can increase the spatial resolution of an image, no interpolation of a single image is capable of increasing the optical resolution of an image. Any perceived increase in detail due to enlargement is a result of limitations of the human eye's visual acuity when viewing the original smaller image.

Both bilinear and bicubic interpolation methods create new pixels that are not in the original image during the enlargement process but they are seen as permissible since the new pixel information is not simply added but rather extrapolated from the original image content in such a way that the content of the captured scene is unchanged. However, interpolation of factor of greater than four is not recommended as this can in fact misrepresent the content of an image as too many pixels in the resulting enlargement are created through extrapolation rather than by the image capture device [18].



(a) (b) (c)

Figure 3.1: Simulated duplicative vs interpolative enlargement [59]
(a)Original, (b)Duplicative Enlargement (Nearest Neighbor), (c) Interpolative enlargement.



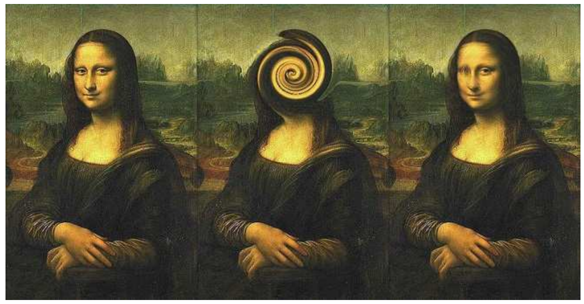
(a) (b) (c)

Figure 3.2: Interpolation comparison
(a)Nearest Neighbor, (c) Bilinear, (c) Bicubic

Distortion correction

Another use for interpolation is in restoring the dimensions of objects in an image should they become distorted. Distortion correction is comparable to stretching an image that is printed on a piece of stretchable rubber, like the surface of a balloon. Distortion correction uses interpolation similar to enlargement techniques but differs from it in that the enlargement does not occur uniformly throughout the image. Distortion correction is a necessary step for photogrammetry as any distortion will decrease the accuracy of image measurements.

Distortion comes in many forms. They could be maliciously inserted into an image, such as the famous case of Christopher Paul Neil who hid his identity by using a swirl distortion on images of his face and was later identified by a de-swirl correction in the opposite direction as in [60] (Fig. 3.3). However, distortions in images are usually naturally occurring.



(a) (b) (c)

Figure 3.3: Distortion correction of a swirl.
The original (a) was swirled (b) and then de-swirled (c) with only minor loss of detail.

The most common cause of distortion is a camera lens. Lens distortion manifests in the curving of features in an image making straight lines appear curved, especially toward the outer edges of an image frame. Camera lens manufacturers have to make many compromises in the production of a lens regarding form factor, cost, weight etc. Due to the difficulty and cost involved in creating distortion free glass, most camera lenses will have some form of noticeable distortion, especially visible near the edges of the image frame. Very wide angle lenses have a tendency to show a "barrel" aberration while very shallow lenses will often display a pincushion effect (Fig. 3.4). Distortion in images and videos can be corrected manually if there are known straight lines that appear curved that can be used as a reference for correction. When no straight lines are present in a scene but the camera which took the evidence image is available, it is often possible to generate parameters for removing the distortion in the evidence image by correcting another image taken with the camera or by using a camera calibration pattern along with software which can determine the nature of the camera lens distortion (Fig. 3.5).

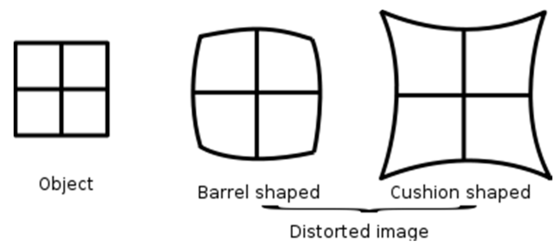


Figure 3.4: Lens distortions

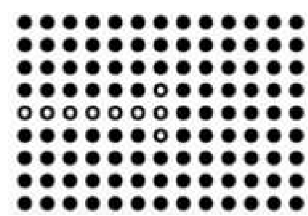


Figure 3.5: Camera calibration pattern

Often, distortion may only be present in one axis of an image. This type of distortion is typical in video recorded in standards such as 2CIF (704 x 288 resolution) where the vertical resolution is halved [43]. Image de-interlacing will also cause a halving of the vertical dimension as half of the image content is removed in horizontal lines along the height of an image. One easily overlooked cause of this type of distortion is the display of an image meant for an analog system with non-square pixels on a digital screen with square pixels. In order to correct these visual errors and give a more accurate depiction of the dimensions of objects in an image, interpolation can be used to enlarge along only one axis of an image.

Lastly, distortion correction can also be helpful in correcting skewing of features caused by perspective. Perspective distortion refers to the optical effect that causes objects farther away from an observer to appear small, and objects to be foreshortened based on the angle of view. Perspective has been used in art for centuries for dynamic effect and to give the illusion of distance in drawings and paintings. In forensics however, perspective distortion can often be an unwelcome effect that misrepresents the dimensions of an object captured in an image, making it difficult to get accurate measurements and hinder visibility of an object. Forensic photographers often go to great lengths to capture images with the camera parallel and perpendicular to the plane of an object being photographed in order to minimize perspective distortion as much as possible. If an image is not photographed in this way, it may be necessary to correct distortion for proper measurement through a process called rectification. Rectification can be used to correct the perspective of an object, given that two conditions are met: 1) The object sits evenly on a flat surface. 2) There is an object of known dimensions on the same plane as the target object. If these two conditions are met, the dimensions of the unknown object can be determined and corrected based on those of the known object [21].

This technique has a couple of limitations. Distortion correction is limited to objects within the known plane and planes that are parallel to the known plane. Objects not parallel to the known plane will become further distorted after the distortion correction process. Furthermore, certain areas of an image will not appear very clear depending on the amount of interpolation needed to stretch the distorted area to the required size.

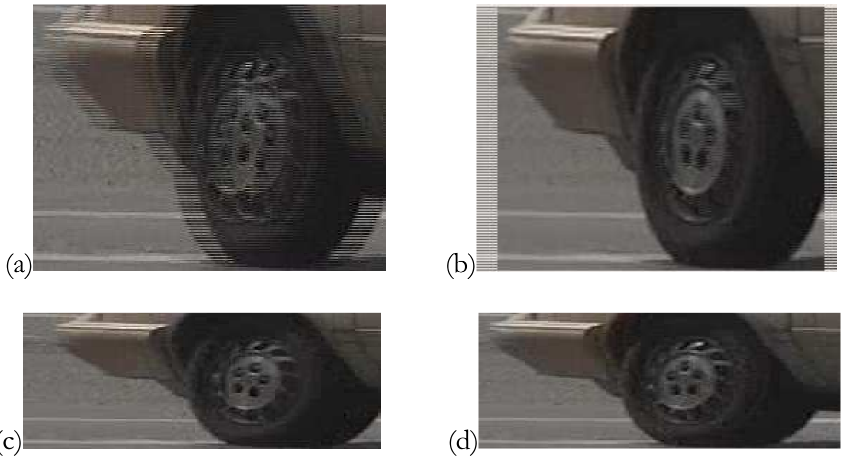
(a) (b)

Figure 3.6: Rectification before (a) and after (b)
Notice that image details are more blurred on the left side of (b) due to the amount of interpolation needed to resize that area.

De-interlacing

Interlaced video works in conjunction with the feature of the human eye known as eye persistence [43]. Eye persistence is what "tricks" the brain into seeing motion in a rapid sequence of still images. If the sequence of images is not fast enough, video will appear to flicker like a quickly moving slide show rather than a steady moving motion. Interlaced video doubles the perceived speed of frame refresh rate in a video by alternating between refreshing one of two interwoven images (fields) rather than refreshing the entire video frame at once. This video technique is a holdover of the bygone analog era where it was commonly used to carry video for the television broadcast standards PAL, SECAM, and NTSC. Interlacing was useful because of its ability to save video broadcast bandwidth and because it was an answer to technological limitations in analog television refresh rates which could not refresh the entire screen fast enough to reproduce the number of pictures per second needed for the human eye to see uninterrupted flicker free motion. Interlaced video is becoming scarcer nowadays but it is still used in High Definition TV broadcasts in the form of 1080i (Interlaced frames of 1920 x 1080 resolution).

Fields in interlaced video are woven in a set of even and odd horizontal line pairs in frames of video in such a way that even lines make up one image and odd lines make up another. Because the fields are refreshed at different times, two very dissimilar fields can appear at once causing difficulty in viewing a scene. This is especially true during scenes with heavy motion or rapid scene changes. De-interlacing is the act of separating the odd and even fields which compose the frame of a video. By de-interlacing a video frame or group of video frames it is possible to improve the overall visibility of a video. Besides simply de-interlacing, some programs allow moving fields from left to right in a process called field alignment, so that the even and odd fields better coincide to form a whole picture.



(a) (b) (c) (d)

Figure 3.7: Interlacing operations
(a) An original interlaced image, (b) Field alignment, (c) Even lines separated, (d) Odd lines separated

De-interlacing should generally be done before any other image enhancements. Because of the impairment of view that interlaced frames cause, it is difficult to tell what type of processing is suitable in an image while it is still interlaced. Furthermore, any processing that changes the relationship between neighboring pixels, such as lens correction, can make it impossible to de-interlace video frames into their correct respective fields. Because interlaced frames loose half of their pixel information (either the even or odd fields) it is necessary to resize the Y axis of a de-interlaced video or video frame in order to show the correct dimensions of objects after de-interlacing.

Color and intensity adjustments

Color and intensity adjustments are among the most frequently used, if not *the* most frequently used, image enhancement techniques. Tonal adjustments are often necessary due to the eye's inability to distinguish between closely related color and intensity values and also due to many cameras' poor tonal reproduction performance in certain conditions such as low or high light. Intensity corrections are useful for an increase in contrast between image features and color correction is used to give an accurate portrayal of the colors in a scene at the time that an image was created or to present images in a neutral color (Fig. 3.8)

(a) (b) (c) (d)

Figure 3.8: The infamous gold-white/black- blue dress
(a) Original image. (b), Color corrected.
(c) Color corrected and contrast adjusted.(d) Retailer image of dress
Image (a) made international headlines in mid-2015 when some claimed it to look white and gold while others claimed it to be blue and black. White balance and tonal adjustments reveal its true color.

Color in digital images is a combination of hue and intensity. Hue is the differentiation between different colors, i.e. orange, green, and burgundy, etc. Intensity can be thought of as the brightness or darkness of an image denoted by where on the tonal range (0 -255 for an 8 bit image) the pixels values sit. While each color channel will have intensity values associated with it for each pixel, intensity itself can be thought of as colorless. To see the relationship between color and intensity in a standard RGB image, it is useful to examine the interaction of color and channels and intensity values at a single pixel. Assuming an 8 bit image, a pixel's color may be a deep purple if it is contains the values Red=150, Green=50, and B=100. If only the green value is increased by 150 intensity levels, the new R=150, G=200, B=100 mixture would result in the pixel becoming green. If instead 100 intensity values were added to all three pixels at the same time, the resulting R=250, G=150, B=200 color combination would result in a lighter shade of purple than the original values. This is the basics of color and intensity in images. Uniformly changing all channels increases the overall intensity of a pixel and non-uniformly modifying the relationships between color channels will change the hue. [61]

Color space

A system for managing the possible colors that a system can reproduce is called the color space. Up to now, we have thought of color as a mixture of Red, Green, and Blue, color channels but RGB color is only one of many color spaces used in digital images. The distinction between intensity and color can become even clearer if a different color space is used such as YCbCr, as mentioned in Chapter 2. The YCbCr color space can be used to encode RGB information in a way that separates color components from intensity components, with Y being luminance and Cb and Cr being blue and red chrominance information respectively. This color space takes into consideration that people are more susceptible to changes in intensity than changes in color. Separating luminance from chrominance by converting an RGB image to the YCbCr color space can be especially useful for making color corrections without affecting the overall intensity of an image. Other color spaces exist such as the subtractive CMYK color space used to in printers using Cyan, Magenta, Yellow, and Black ink. This color space is not often used for working with digital images unless they are meant to be printed after processing is complete.

White balance

White balance is a form of color correction for adjusting hues in an image so that they are not oversaturated by a particular hue. Hues may be skewed toward a particular color in an image depending on different lighting sources such as incandescent lights, fluorescent lights, an overcast sky, etc. If the hues are well balanced, colors like white and gray will appear to be a neutral tint. Color adjustments in forensic images are not based on aesthetics so white balance adjustment must be guided by a reference denoting what a neutral white, black , or gray should look like. White balance is a global adjustment which typically affects every pixel in an image in the same way. Thus a white balance correction that makes an image like Figure 3.9 (b) look like Figure 3.9 (a) could involve uniformly subtracting 50 red intensity values from every pixel in the image.

(a) (b)

Figure 3.9: White Balance Correction.

<u>The histogram</u>

An image histogram is a powerful tool to not only visualize the intensity and tonal values of an image but to adjust them as well. The histogram is a type of bar graph which displays how tonal values are distributed among the pixels of an image. This is organized with the number of pixels on the Y axis and the tonal range on the X axis. By viewing a histogram, one can quickly discern where pixel information lies and what can be done to improve its visibility. In most image processing programs, a histogram will have a set of high and low bounds which can be adjusted to redistribute intensity values found within. For example, if an image has a dynamic range of 0-255 but most of the pixels in the image are in the mid-level intensity values between 80 and 150, the histogram's lower bound could be adjusted to 80 and the high bound could be adjusted to 150. Any tonal value 80 and below would then be considered a 0 intensity value and any tonal value 150 and above would be displayed as a 255 tonal value. Afterwards the remaining values within these bounds would be stretched within the 0-255 range. This creates a larger disparity between closely related intensity values than that found in the original image which results in an increase in contrast. When the histogram is used in this way, it does not create new information but rather changes the way that the already existing information is presented. The downside to this method is that information outside of these bounds is fully clipped. If the clipped information is irrelevant to an investigation, then this may not be a problem, but if the information is relevant, then other enhancement techniques will be needed.

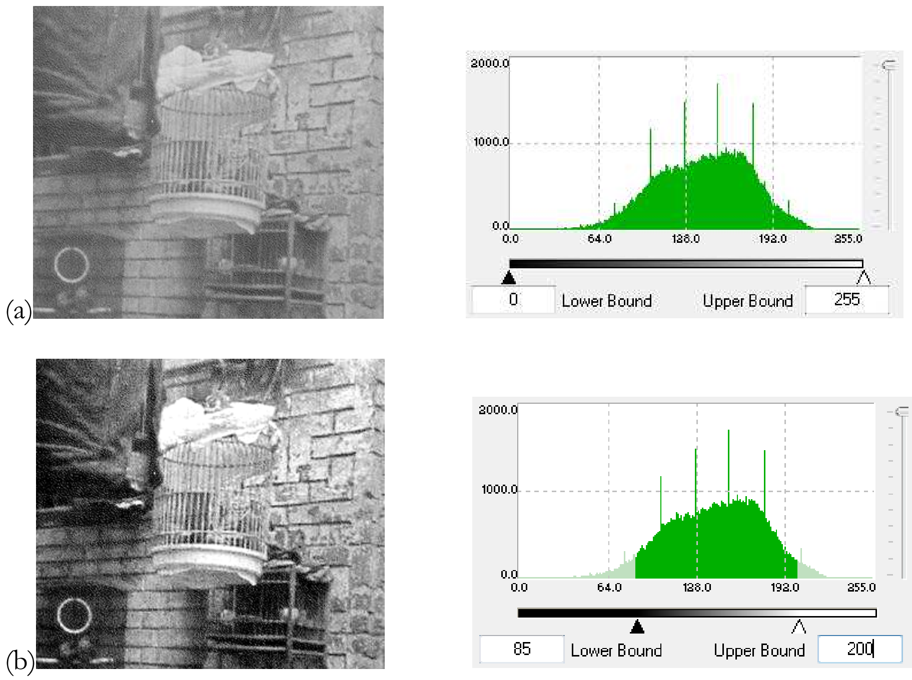


Figure 3.10: Histogram bound adjustment.
Notice that some details in the cloth above the birdcage are clipped in (b) as the histogram bounds were adjusted.

<u>Histogram equalization</u>

Histogram equalization is a standard technique for redistributing the global tonal values in an image in such a way that they are more equally represented throughout the image. The spread of pixel values in the histogram equalized image results in a flatter shape of the image histogram. This method of histogram adjustment is an improvement over moving histogram bounds because much less information becomes clipped. However, there are two main drawbacks of this method. Since global histogram equalization processes all areas of an image equally, the process may not preserve details well [62]. Also, it may bring out the appearance of noise in an image as the intensities of the pixels that make up noise are used to calculate the overall tone equalization levels [62].

Another method for achieving a similar result is local histogram equalization. Local equalization modifies tonal values of individual pixels with respect to their surrounding neighboring pixels as opposed to global equalization which takes the entire tonal gamut in the image into account. Local equalization can be weighted, meaning that the pixel to be modified at the center of a group of pixels will influence the equalized result more than pixels farther away [21]. Local equalization effectively flattens the tonal ranges of images, creating very little contrast except at edges. This can be good for observing detail in an image but not for overall visualization as this process creates relationships between tonal values that are unnatural and misleading.

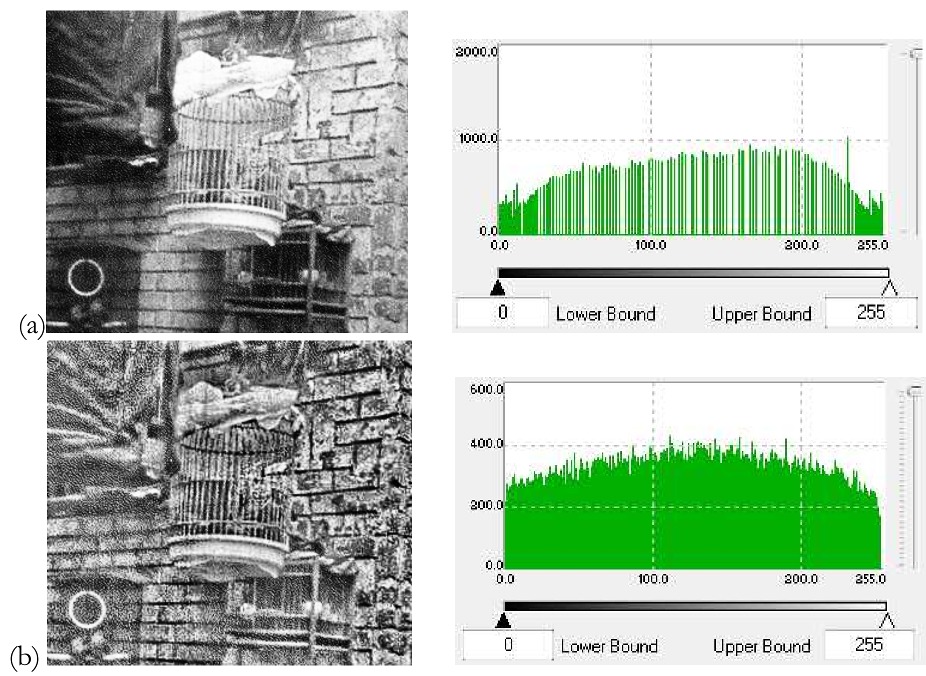


Figure 3.11: Global histogram EQ (a) and local histogram EQ(b)
Notice that all the detail that is found in the global EQ is also available in the unprocessed image, only now it is more apparent.

Noise reduction

Image noise refers to variations or disturbances in brightness or color information in an image that do not reflect the actual content of a photographed scene [12]. These disturbances manifest as what can be described as small spots of color, dust, or snow superimposed on an image. Noise can come from many sources. The appropriate techniques used to reduce noise often depend on the source and type of noise.

Sources of noise

Random noise can be introduced at many stages of the image creation pipeline. Both CCD and CMOS sensors produce random electronic noise through the different ways that electrons operate within them. In a CCD censor, electrons pass through the sensor at one end, are stored depending on the limits of a capacitor, and are discharged at the other end. The storage and movement of electrons from one end of a sensor to another causes CCD noise to be great in high intensity regions. Meanwhile, in CMOS chips electrons are discharged independently by each individual light detector resulting in more noise in low light areas [18]. Random noise is especially apparent at high ISO settings which determine camera sensors sensitivity. In addition, random noise can be the result of thermally generated electrons interacting with the imaging sensor [43]. This noise is ever present at temperatures above absolute zero [43]. Other causes of random noise can occur during the image processing stage due to analog to digital quantization errors, errors in transferring bits to the storage medium, and demosaicing artifacts as discussed in Chapter 2.

Noise can also be non-random. In both CCD and CMOS sensors, the individual light sensitive diodes may be more or less sensitive resulting in fixed pattern noise that is unique to a camera. Fixed pattern noise is often not as discernable as random type noises and does not usually degrade an image significantly. In fact, rather than there being a need to reduce this noise, noise reduction techniques are commonly used to isolate these "hot pixels" from non-fixed pattern noise for forensic camera identification purposes. If there are defective pixels in a sensor, another type of fixed noise will appear which manifests as completely black or white specs in what is known as Salt and Pepper noise.

Nonrandom noise is generally easier to reduce than random noise because it is repeatable [63]. If the pattern of the noise can be determined, the noise can simply be subtracted from an image. In fact, many CMOS based cameras do just that. A function is built into the sensor that captures a dark image quickly followed by the intended shot, at which point the dark image containing the thermal, pattern, and/or salt and pepper noise can be subtracted [43].

Signal to noise

The histogram can be a great tool for visualizing noise and understanding signal to noise ratios. Signal to noise describes the ratio between the information in an image that corresponds to the actual scene content and the information that is a result of noise. If a histogram of a solid gray image was taken, there would be one long bar in the histogram demonstrating that all the pixels in the image have the same tonal value. This is the highest signal to noise ratio possible. By adding noise to this gray image, new bars appear in the histogram corresponding to the amount and type of noise added. As more and more noise is added to the image, the original bar of gray will shrink until it is no more. This is the lowest signal to noise ratio possible.

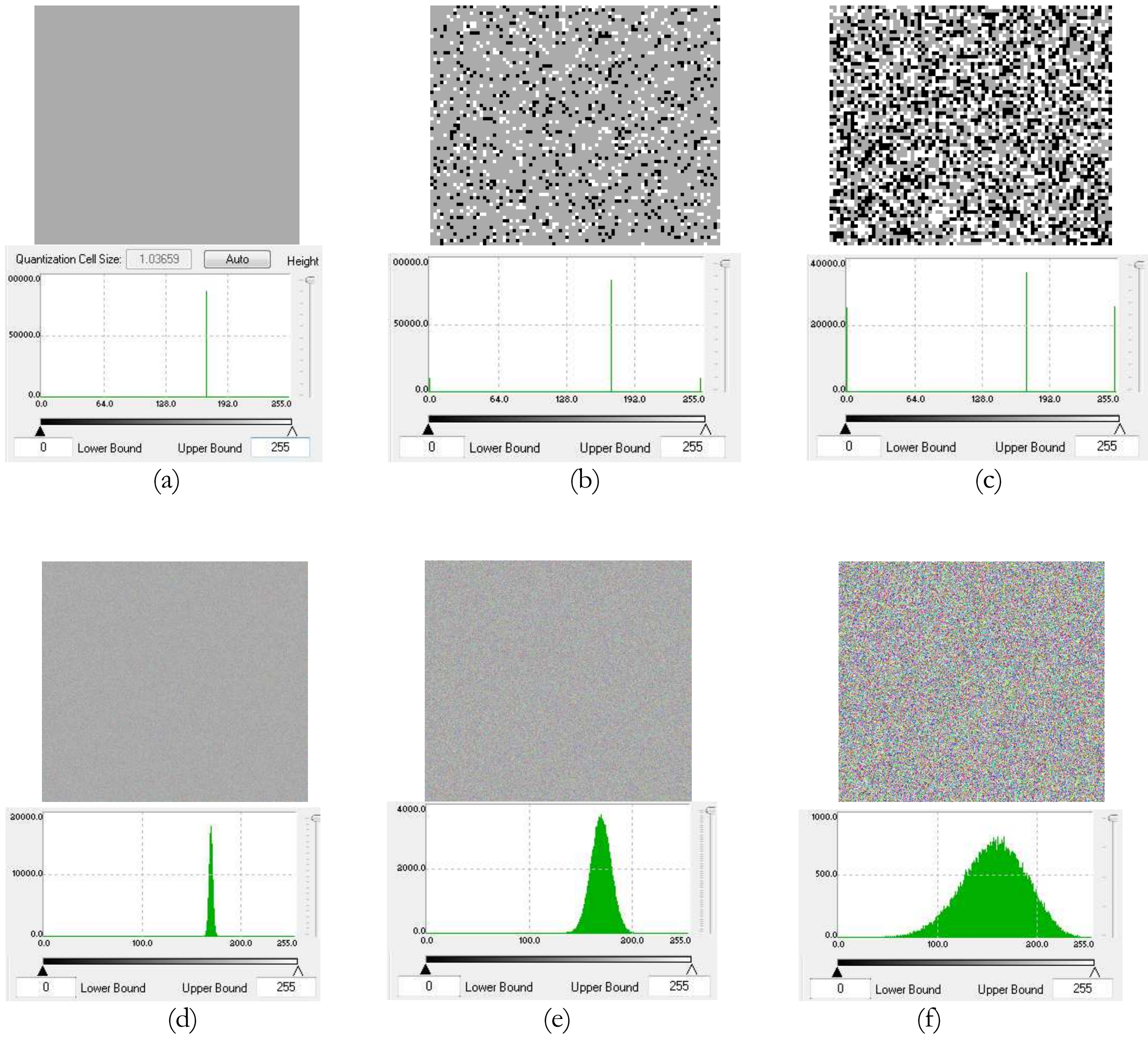


Figure 3.12: Signal to noise
(a) Original image signal, Highest S/N, (b)Salt & Pepper Noise (c),Very low S/N S&P Noise
(d)Slight color noise added, (e) Medium color noise, (c) Very low S/N, high color noise

<u>Noise channel filtering</u>

When working with an RGB image, it is often beneficial to noise reduce the blue color channel independently of the others. Noise often manifests strongly in the blue color because sensor chips in cameras are often less sensitive to blue light [21]. Many noise removal filters have a tendency to blur edges or remove fine details. By working in the blue color channel only, the fine details are preserved in the other two color channels.

By the same token, many images can benefit from noise reduction of intensity only. To do this in an RGB image, it must first be converted to YUV or YCbCr color so that the Y channel, the luminance, can be independently filtered.

Spatial domain techniques

A common technique for removing noise in the spatial domain is spatial averaging, where pixel values are mathematically averaged with a Gaussian function to create a similar distribution of values. This works well for reducing the appearance of noise in uniform areas of an image but does not work well in non-uniform locations such as edges. Near edges, averaging combines very dissimilar values together resulting which results in blurry edges. Though special rules can be made so that areas containing edges contribute less to the overall average, there is no good way to diminishing this blur effect completely as of yet [21].

If the appearance of blur is a concern after denoising, the median filter can be used. The median filter is a standard filter for noise removal in the spatial domain. This filter analyzes every pixel in an image in relation to its surrounding neighborhood of pixels and ranks their values in relation to their intensity. The value at the middle of this rank will then become the new value for the center pixel of the neighborhood. This filter works on the assumption that noise will be either darker or lighter than the median pixel value for a given neighborhood of pixels. Of course, this assumption will not always hold and generous application of this filter can remove fine details from an image. The median filter results in a denoised image that is composed solely of intensity values that were already found in the original image. This is a popular noise removal method due to its speed and because it does not blur edges, unlike other noise reducing techniques [21].

The extrema filter is a noise removal method for removing pixels at the extreme low and high tonal ranges of a given pixel area. It gives results that are similar to the median filter but without ranking pixel values. This filter works best on salt and pepper noise as this noise is at the highest extremes of intensity value. Much like the median filter, new pixel values for the extrema are taken from the pixel values neighboring the extrema.

Though the appearance of noise can be mitigated, image information is invariably lost irrecoverably due to noise (Fig. 3.13).

(a) (b)

Figure 3.13: Noise removal
(a)Original, (b) Noise added globally and removed on a local area.
Noise in b is removed but the original sharpness and color of the image is diminished.

Frequency filters

By converting an image to the frequency domain, bandpass filters can be used to attenuate noise. Noise corresponds with high frequencies of an image in the frequency domain. A low pass filter blocks high frequencies and allows lower frequencies to pass undisturbed. Once the high frequencies are attenuated or removed, the image is returned to the spatial domain resulting in a noise reduced image. This technique, like all denoising techniques we've seen so far, has the potential to reduce non-noise details.

If noise only manifests within a specific range of frequencies, a band stop filter could be used for noise removal at that specific frequency range. This could potentially preserve small details that are above or below the frequency range corresponding to the noise. Stop band filters can be very useful for periodic image features and noise such as marks from a halftone printing process and patterns on an image. Periodic noise can be displayed in a graphical representation of an image's Fast Fourier Transform [18][21]. In the FFT, periodic noise appears as points of intensity. The farther a point is from the center denotes a higher frequency. Therefore the central noise consists of the lowest frequencies which correspond to the non-noise portions of the image. As seen in Figure 3.14, by using a stop band filter on these stars (except the central one), reoccurring noise and patterns can be greatly diminished in an image.

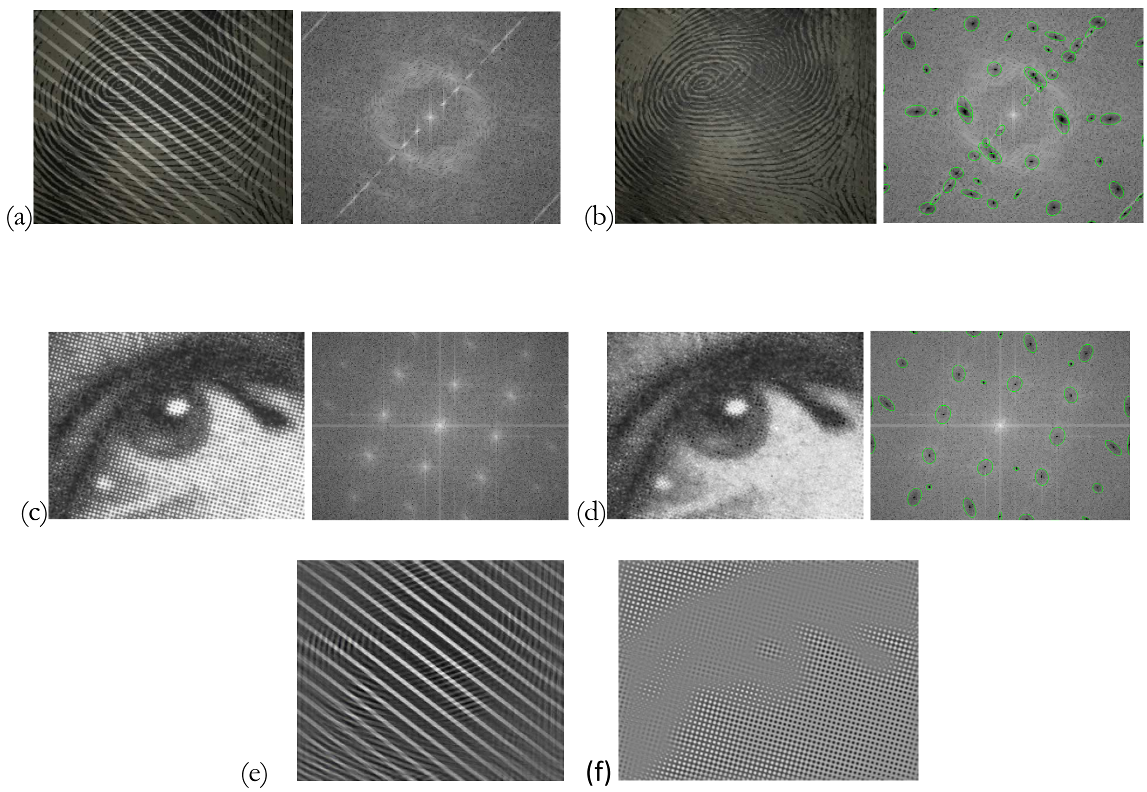


Figure 3.14: Band filtering with a Fourier transform
(a)Original patterned image and 2DFFT, (b)Filtered image. (Band Stopped Selection in Green)
(c)Original halftone image and 2DFFT, (d)Filtered image. (Band Stopped Selection in Green)
(e) The noise filtered from a to generate b, (f) The noise filtered from c to generate d

Sharpening and deblurring

Sharpening and deblurring are generally thought of as two separate processes but they share many commonalities. Both can be used to make distinctions in edges (areas of high gradient magnitude) and fine details and can be carried out in either the frequency domain or spatial domain through similar means. Edges and other image content can become convolved as pixel information corresponding to an object becomes diffused throughout an image, as illustrated in the pinhole camera model in Chapter 2. In digital images, this can be caused by compression, image processing artifacts, camera lens settings, and movement of either the camera or the scene during the creation of an image. Sharpening and deblurring can be used to correct these effects to different extents.

Sharpening

Just as a low pass filter has been shown to reduce high frequency information like noise, a high pass filter can increase the visibility of edges and fine details. Sharpening can also be accomplished in the spatial domain with a classical tool called an unsharp mask. The unsharp mask dates back to the time of film, though it can be much more easily applied in the digital age. In this method, a slightly blurred version of the original image is generated and then subtracted from the original image. Because the blurred image contains low frequency information, what is left after subtraction is high frequency information. The high frequency information can be added back to the original image to create a sharpened version of the original. The unsharp mask can be mathematically demonstrated to be equivalent to its high pass filter sibling, and is therefore accomplished in the frequency domain in most image processing software rather than elaborately adding and subtracting images from each other.

Alternatively, sharpening can be carried out with the use of an image kernel. A kernel is a matrix of numbers which can be used to generate a new pixel value based on a weighted average of a neighborhood of pixels. Depending on the values that make up the kernel matrix and the overall matrix size of the kernel, different effects can be achieved such as blur, edge detection, and sharpening. The numbers in a kernel correspond to the weight, or importance, that a neighboring pixel's value should play in affecting the central pixel [64].The application of a kernel is described in the following steps:

1) The kernel is centered over a pixel in the source image.

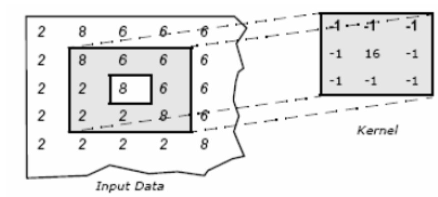
2) The overlapping values of the kernel matrix and source image pixels are multiplied.

3) The resulting values from multiplication are added together.

4) The sum of these values is divided by the sum of the kernel

5) The central pixel in the source image is given the value of the quotient of step 4 in a new image.

6) This is repeated for all pixels in the source image.



[(-1*8)+(-1*6)+(-1*6)+(-1*2)+(16*8)+(-1*6)+(-1*2)+(-1*2)+(-1*8)] ÷ (-1-1-1-1+16-1-1-1-1)

New central pixel value = 11

Figure 3.15: A 3x3 image kernel [64]
The kernel is defined as [-1-1-1;-1-16-1;-1-1-1].

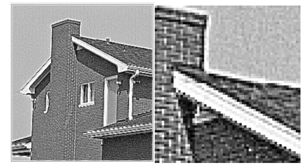
Regardless of the sharpening method used for sharpening, the overall effect of sharpening will be to heighten the magnitude of difference between edges. Over sharpening can create unnatural portrayals of fine details with a common artifact being halos of brightness around edges [65] as can be seen above the roof in Figure 3.16(c).

0	0	0
0	1	0
0	0	0

(a)

1	1	1
1	1	1
1	1	1

(b)



0	-1	0
-1	5	-1
0	-1	0

(c)

Figure 3.16: The effect of different kernels [66]
(a) No change, (b) A blur kernel, (c) A sharpening Kernel

<u>Deblurring</u>

The idea of deblurring is thought of by many as being as realistic as descrambling an egg. Luckily, deblurring is indeed possible and the basic theory needed to accomplish it can be easily expressed with simple algebra. When objects are blurred in an image, their color information is diffused from a central point. The direction and magnitude of this spread of information from a central point is known as the Point Spread Function (PSF) [67]. A blurred image can be described with the equation $g=Hf + n$, where g is the blurred image, H is the (PSF), f is the theoretical image had it not been blurred, and n is noise added when the image was created [68]. If the point spread function can be known or calculated, an image can then be deblurred by solving for f and estimating the signal to noise ratio for the image.

Calculating the PSF of a particular blur type can be difficult but it is possible. To see how this may be done, consider a starry night sky. A star in the sky free of blur should theoretically appear as a single dot in an image but due to optical imperfections in a camera it will not. If an image kernel can be estimated that will blur a single dot to the same degree as the star, then the inverse of the kernel can be applied to the entire image to undo the blur [21]. This technique of correcting blur was used with the Hubble telescope before its optics where corrected by the installation of a secondary mirror [21].

The method described above for calculating a PSF and deblurring an image is known as non-blind deconvolution, because the PSF can be known. When a PSF cannot be accurately calculated as above, it may be estimated and an image can be deblurred through what is referred to as blind deconvolution [68]. Blind methods of estimating PSF may include feeding a deblurring algorithm a general estimation of the length and direction of a blur or tracing the path of a blur with computerized tools so that the algorithm can understand the direction, at which point the magnitude of the blur can be estimated heuristically. Other methods involve using series of clean and blurry images that a computer algorithm can use to estimate what PSF type is needed to get from one image to another.

(a)

(b)

Figure 3.17: Lens blur restoration
(a) Original non-synthetic blur (b)Blur reduced

(a) (b)

Figure 3.18: Motion blur restoration [68]
(a) Image with motion blur, (b) Image a deblurred. Result has ringing artifacts in the corners

Finding the exact PSF for a particular image is not always possible. In fact it may sometimes be impossible to find as different points in an image will be blurred to different extents based on their distance from the camera and the type of motion occurring either in the scene or by the camera itself. Deblurring requires a high resolution image for results such as those seen in Figures 3.17 and 3.18. If deblur settings are improperly adjusted for an image, halos around prominent edges may appear known as ringing artifacts.

Video enhancement techniques

All of the previously demonstrated image enhancement techniques can be used on both single images and videos on a frame by frame basis. Conversely, the following techniques cannot be performed on single images as they are realized by extrapolating information from multiple frames of video. In some cases, performing these enhancements on multiple still images rather than frames of a video is possible. However, in order to generate consistent results, it is often required that the images be taken with the same camera, at the same location, and within a small time period from each other. Frames of videos automatically satisfy these requirements.

Frame averaging

Frame averaging, also known as temporal averaging, is a noise reduction technique in which the pixel values of multiple images are averaged together, resulting in a higher signal to noise ratio. The effect of this technique is equivalent to setting a long exposure on a camera. Therefore frame averaging can have the associated consequences of long exposures. Frame averaging can only reduce random noise because it will be different at every frame, causing it to contribute less and less to the overall average in relation to the actual scene contents which are constant at every frame. This results in the random noise being nullified after averaging. This also means that nonrandom noise such as fixed pattern noise will be preserved. For this reason, frame averaging is often used to generate fixed pattern noise "finger prints" for camera identification. Similar to setting a long exposure, any movement of the camera or objects in a scene will create blur, since pixel information for one object will be represented at multiple locations of an image. Therefore this technique is only possible if the scene content is still, or if the object of interest in the video is stabilized prior to averaging so that it remains in the same location along every frame of video.

10	10	10
8	8	8
2	4	2

(a)

4	4	4
6	6	6
8	8	8

(b)

7	7	7
7	7	7
5	6	5

(c)

Figure 3.19: Averaged values in two 3x3 pixel grids
(c) is the average of (a) and (b)

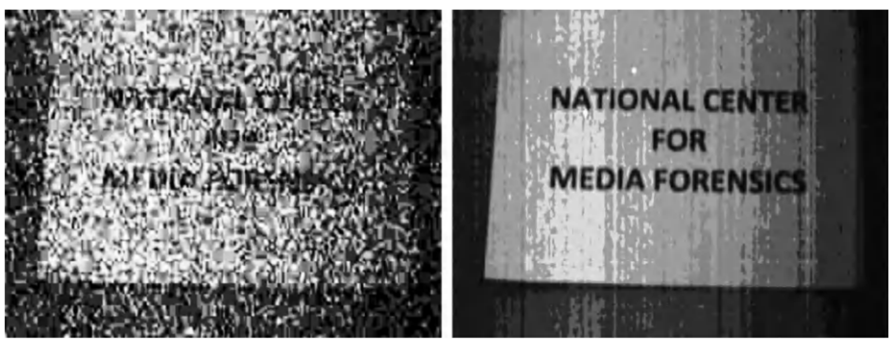


Figure 3.20: Example of frame averaging [15]
The left side is the example of one noisy frame. The right side is the result of several averaged frames.

<u>Super resolution</u>

Super resolution is a natural successor of frame averaging. Super-resolution is a technique that is useful not only for denoising but also for creating an image with high spatial and optical resolution by extrapolating information from a series of images with lower spatial and optical resolution. As discussed previously, interpolation of a single image cannot add detail to an image.

However by using multiple images, super resolution techniques can improve image quality beyond what a camera's sensor, optics, and internal electronics is capable of recording.

Super resolution is made possible by observing sub-pixel shifts between images. Because imaging sensors can only record information by whole pixels by definition, super resolution algorithms subsample these pixel values so that minute changes can be identified between frames of video. If the pixel shifts between frames are known or can be estimated, then super resolution is possible. The known movement of pixels can be determined if the coordinate of a camera is known, such as in satellite imagery, or through mathematical estimates. If there is no displacement of pixels between frames or if the displacement occurs in whole integer values, this would mean that the frames will not contain sub-pixel information and super resolution is not possible [69]. For super resolution to be successful, displacement of pixels must occur in such a way that information from one frame cannot be obtained by another.

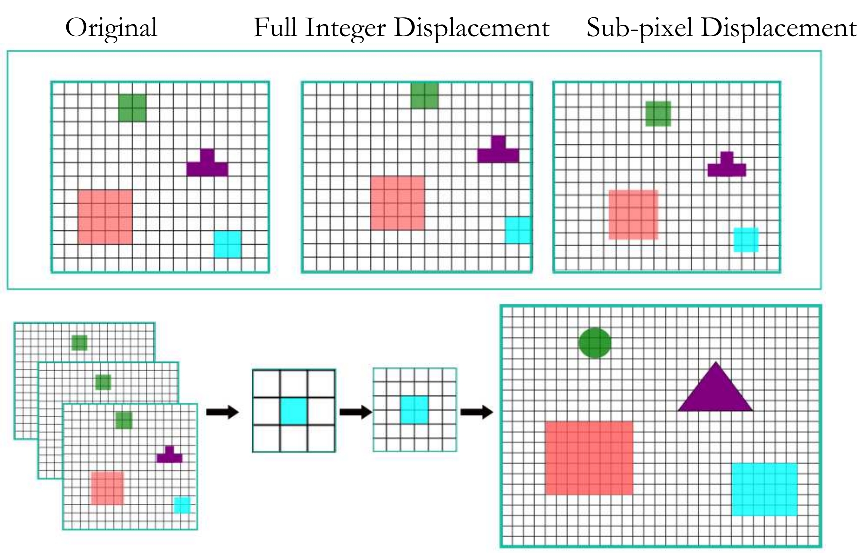


Figure 3.21: Super resolution model

(a) (b)

Figure 3.22: Super resolution before(a) and after (b) [15]

<u>Stabilization</u>

Stabilization is a process for adjusting individual frames of video so that an object or person of interest remains in the same location in the overall video frame during playback. It can be useful for improved visual analysis of a video and as preparation for frame averaging or super resolution techniques.

CHAPTER IV

PROPOSED FRAMEWORK

The main purpose of an order of operations is quality assurance. An order of operations establishes protocols that mitigate human fallibility and promote analyses that are reliable. By integrating an order of operations for image enhancement into an SOP, the QMS can be strengthened overall. If the order is sound, there can be increased confidence that examiners will generate visually comparable results and that these results are less likely to misrepresent the visual data. Furthermore, by writing down expectations in an SOP, documentation exists to can help keep examiners accountable and establish reliability in court.

The proposed framework for digital image enhancement presented ahead concentrates on an order of operations for images which require multiple concurrent image processing techniques. Not every image that a forensic examiner will come across will require all of these techniques to be utilized at once, and in many cases it would not be possible or desirable to apply specific enhancements depending on the type of image file and/or imaged content. In such cases, a step in the proposed order may be skipped and the next step can be continued.

Table 4.1 provides the proposed image enhancement framework. It is based on a modified order of operations developed by the National Center for Media Forensics in Denver Colorado [70], the contents of Chapters 1, 2, and 3 of this work, and research and testing of various image filter combinations. All images processing tests were conducted with the same filter settings, while only changing the order of enhancements. The results of the tests used for the development of his framework are given in the next section along with an expanded explanation of the rationale behind this order.

Forensic Image Processing Framework	
Preparation	
1	Record Log
2	Ensure playback
3	Adjust Bit Depth -> 16 bit
Analysis	
4	Visual Analysis • Assess problems o Noise o Error o Distortion
5	Considerations • Purpose of enhancement • Lab/Personnel capacity
Image Processing	
6	Correct Video Artifacts De-interlace Stabilization
7	Resize*
8	Super-Resolution*
9	De-noise • Frame Averaging* • Pattern removal • Median Filter • Gaussian blur • JPEG artifact removal
10	Deblur • Lens blur • Motion Blur
11	Distortion Correction • Lens correction • Rectification
12	Contrast/Color • Brightness/Contrast • Histogram • Dynamic range reduction • Equalization • Separate color channels
13	Sharpen
Output	
14	File Output
15	Documentation
* Items marked with an asterisk may be completed at the end of all other image processing or in the order listed.	

Table 4.1 Proposed Order of Image Enhancement

Rationale

<u>Acquisition and preparation</u>

Before image enhancement can begin, a digital image must be acquired and stored in accordance with the principles of media forensics. The extent that any image can be relied on, including the reliability of any subsequent enhancements, is dependent on whether it was acquired correctly by well trained personnel who followed forensic best practices. This includes the duplication of files, file hashing, storage, and documentation of chain of custody and processes that an image undergoes before enhancement. As acquisition of digital images is slightly beyond the scope of this paper, discussion of that topic is best deferred to SWGIT, particularly "Section 13 Best Practices for Maintaining the Integrity of Digital Images and Digital Video" and "Section 24 Best Practices for the Retrieval of Digital Video".

1) After an image has been properly acquired, it must be prepared for analysis and processing. This can include enabling a history log. Many image processing software, especially those specifically designed for forensics, will have a record log function that keeps track of all processes in an image. This information can often be saved in the metadata of an image file or as a separate log file. By utilizing a log, documentation of all enhancements can be preserved in order to help satisfy reproducibility requirements for science and the courts. If a history log is not available, documentation can be carried out in writing, by the use of screenshots of parameter windows, and if the program allows, by applying each image enhancement to an image in individual layers which are named with the parameter settings used for each filter.

2) It is not uncommon to have difficulty in displaying forensic video and image evidence. Proprietary video formats from CCTVs are plentiful and there may be difficulty playing back video files even with accompanying proprietary playback software. Part of the preparation process may involve re-encoding video or image data into a format that is better suited to analysis and processing.

Forensic image enhancement best practices indicate that this no pertinent information should be lost during the re-encoding [1]. To this end, lossless encoding is recommended.

3) Adjust the image or video processing program to work with images in 16 bits per channel mode. This may or may not be possible depending on the program used. As seen in figure 2.15, a low bit depth can cause colors to become quantized with higher degrees of error. Working with an image that was originally captured at 8 bits per channel will not make it look better, but it will mitigate quantization error during image enhancement [71].

Analysis

4) Once preparation is complete, analysis can begin, starting with visual analysis. Visual analysis is necessary to determine what enhancements will be necessary and if enhancement is possible or desirable.

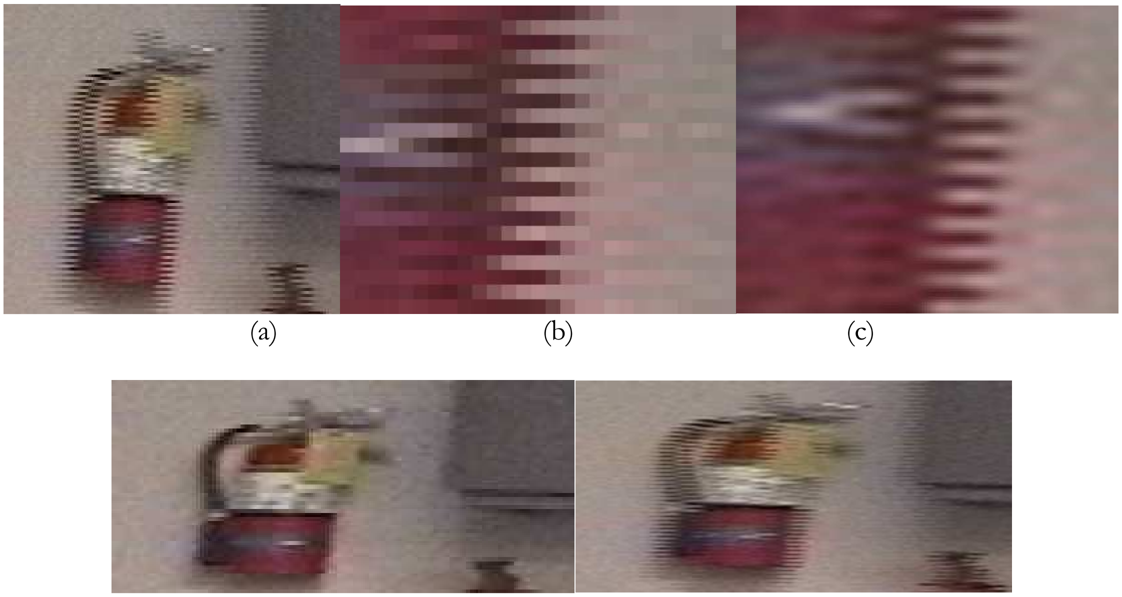
5) The possibility and desirability of enhancement should be considered based on the purpose that the image will serve, the properties of the image, the training of personnel, and lab resources. Often, enhancement may not be necessary at all if the relevant information is apparent without the need of enhancement. Other times, an image may be so badly degraded that qualified personnel will be able to determine on sight if enhancement would be possible. In such cases, enhancement would merely take up laboratory and personnel resources. Personnel in charge of enhancement should be aware of the limitations of their techniques, training, and lab assets including software, hardware, and time.

Before moving on the enhancement part of this framework, there are two important considerations. Image enhancements for forensics are generally made globally. This is done so that the enhanced areas do not misrepresent the image data. Making an entire image brighter for example, may allow one to see dark areas of an image better, but making a small area of an image brighter may make it appear as if the lighting conditions were different than in reality. Making a *very* small area brighter is equivalent to drawing in new pixels and should not be done [18]. The enhancements presented here are expected to be performed globally to all pixels unless it can be made obvious that an adjustment was made on a local area. Also, when it comes to image enhancement, less is more. Strong application of any of these enhancement filters may do more harm than good by introducing harsh artifacts.

Order of operations

6) Correct Video Artifacts:

Interlaced video frames contain two separate images within one frame and therefore cannot be properly analyzed or used with other enhancements until the fields are separated. Interlaced images cannot be accurately de-interlaced after image enlargement, especially after bilinear or bicubic interpolation. Clear distinctions between pixels can be lost after interpolations as distinctive borders between fields are changed or blurred as a result of interpolating content of one field with that of another. Because the two fields are actually two separate images, the interpolation is inherently flawed. Also, de-interlacing algorithms commonly work by separating even and odd single pixel horizontal lines. This cannot be done after enlargement as individual lines of pixels corresponding to a field will consequently be composed of multiple pixel lines. Additionally, processes like lens distortion correction and blur can alter the relationships between pixels, making de-interlacing impossible later on.



(a) (b) (c)

(d) (e)

Figure 4.1: Interaction of de-interlacing and interpolation
(a) Original, (b) Original zoomed section (c) Enlargement with bicubic interpolation
(d) De-interlaced original (e) "De-interlaced" after bicubic interpolation
Figure 4.1 e failed to de-interlace correctly as a result of interpolation.

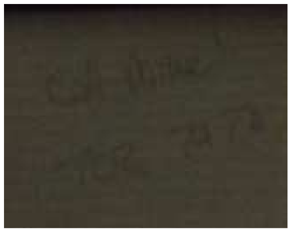
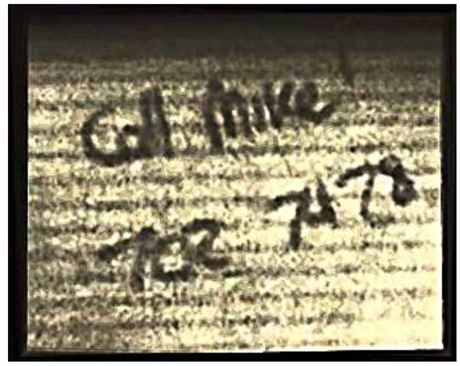
Stabilization is only possible in videos and may not be required before frame averaging or super resolution if no movement is observed within the movie frame. It is listed as a preliminary step as it may help in assessing the content of an image and the determination of future enhancements.

7) Resize:

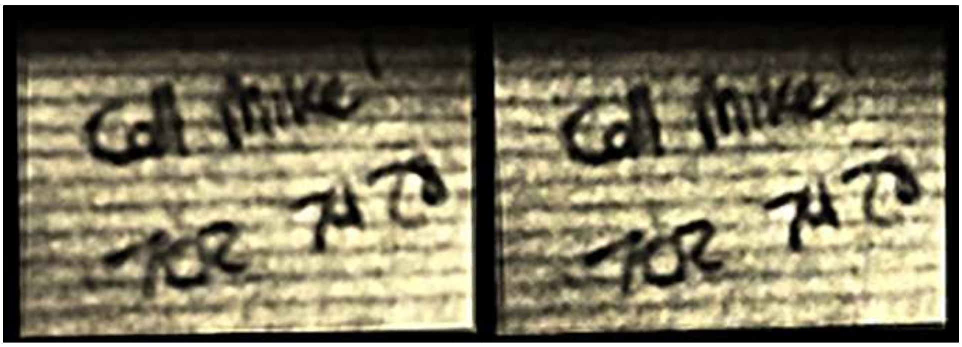
Enlarging an image has both benefits and drawbacks. On one side, the extra pixels resulting from interpolation can add increased precision for future processing of an image. Resizing may also be an aid in noticing features that may be too small to see otherwise in order to better determine a course of action for future enhancement. However enlargement causes two main complications. It increases the overall processing time that other enhancements will require and it can increase the size of unwanted artifacts such as noise. With a modern computer, processing time may not be a concern for a small image, but it can quickly become an issue when attempting to apply enhancements to large portions of video at 30 plus frames per second. The increased size could become a problem for very noisy images. Noise that may originally consist of one pixel in size will increase in size and become diffused throughout the enlarged image since new pixels are interpolated based on noise pixels. This could be countered with a stronger denoising parameter setting later on, but how effective this may be will depend on the nature of the noise after enlargement. Because of this balance between precision, visibility, time, and artifact enlargement, when to resize should be carefully considered. If time and noise are no issue, it is advised to resize as the second step of this order of operations. Alternatively, resizing may be applied after all other image processing has finished.

8) Frame Averaging/Super Resolution

Frame averaging and super resolution are asterisked (*) in the proposed order of operations to signify that they may either be inserted as steps 2 and 3, or applied after all processing is complete. Early processing with these techniques will not allow for individual frames to be processed later on, except with a separate copy of the video data. The result of applying these techniques at the beginning or end of the image processing could aid other enhancements due decrease in noise and, in the case super resolution, an increase in spatial and optical resolution. Alternatively, completing all image processing on individual frames before super resolution may aid the super resolution algorithm better determine how sub-pixel information should be fused in the super frame.

(a) (b)



(c) (d)

Figure 4.2: Effect of deviations in the proposed order of operations
(a) Original image from a video, (b) Original image processed with image processing framework in numerical order (deblur and rectification were not necessary),
(c) Original image processed with the same settings as figure b, without super resolution. Frame averaging was done as the final step. (Original image processed with the same settings as figure b, without frame averaging. Super resolution was done as the final step.

9) Denoise:

Noise is one of the greatest factors that determine the result of an image enhancement. Enhancement filters have no way of making a distinction between what is noise and what is a desired image signal, causing noise to interfere with the processing of an image. Enhancement processes such as histogram equalization and contrast adjustments will enhance the noise by heightening the disparity between noise tonal values and the desired image signal. Noise is particularly damaging during deblurring processes. A blurred image consists primarily of low frequency information. During the deblur process, high frequency information is amplified in order to restore details in an image. Any noise in the blurred image will consequently be amplified [72]. Therefore, it is recommended to denoise an image as the third step of this order of operations, before deblurring. At this point, it may be warranted to denoise a video with frame averaging. Before doing so, one should consider that frame averaging will merge all averaged frames into one, resulting in a static image that can no longer be viewed as a movie. The resulting still image could not be used in super resolution techniques.

The examples in Figures 4.3 and 4.4 below were deblurred by using a Wiener filter at the same settings for every image. The Wiener filter takes into consideration the estimated signal to noise ratio of an image as a user input parameter. The highest signal to noise ratio possible within the software program used was set in order to amplify the visibility of the resulting noise. The test data in figure 4.3 (d) is a 50 frame movie consisting of image b with random noise at every frame. The signal to noise ratio was lower than that of image d. Figure 4.3 (d) is a single deblurred frame from this movie. Image h was created by first denoising the 50 frames used for g with frame averaging and then deblurring the image.

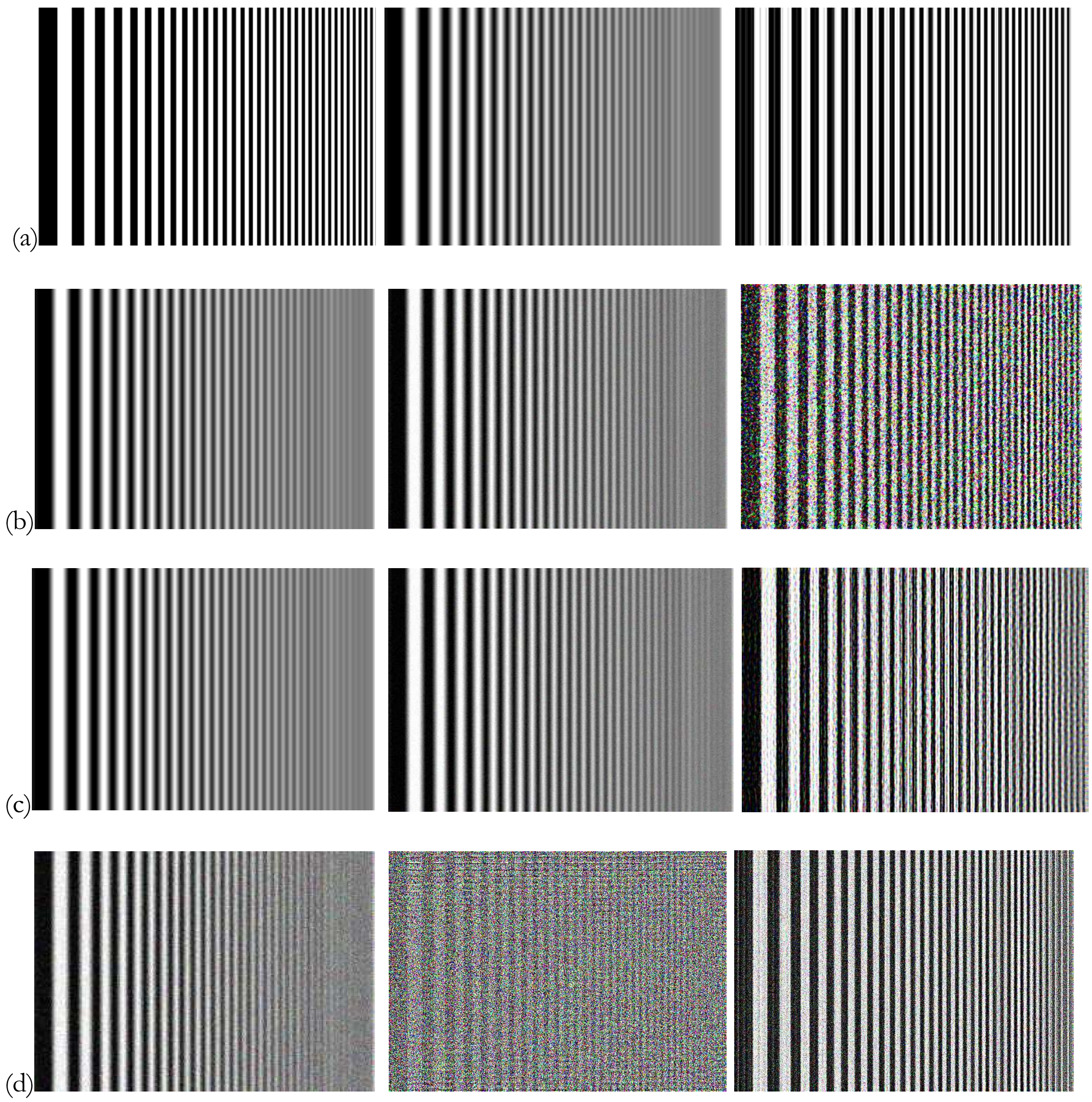


Figure 4.3: Interaction of denoise and deblur
(a) Original image > original with 10 pixel synthetic blur > deblurred
(b)Original with added blur -> blurred with noise -> deblurred then denoised with median filter
(c)Original with added blur -> blurred with noise -> denoised with median filter then deblurred
(d) Blurred and noisy movie -> Single frame deblurred then denoised -> original blurred input first frame averaged then deblurred

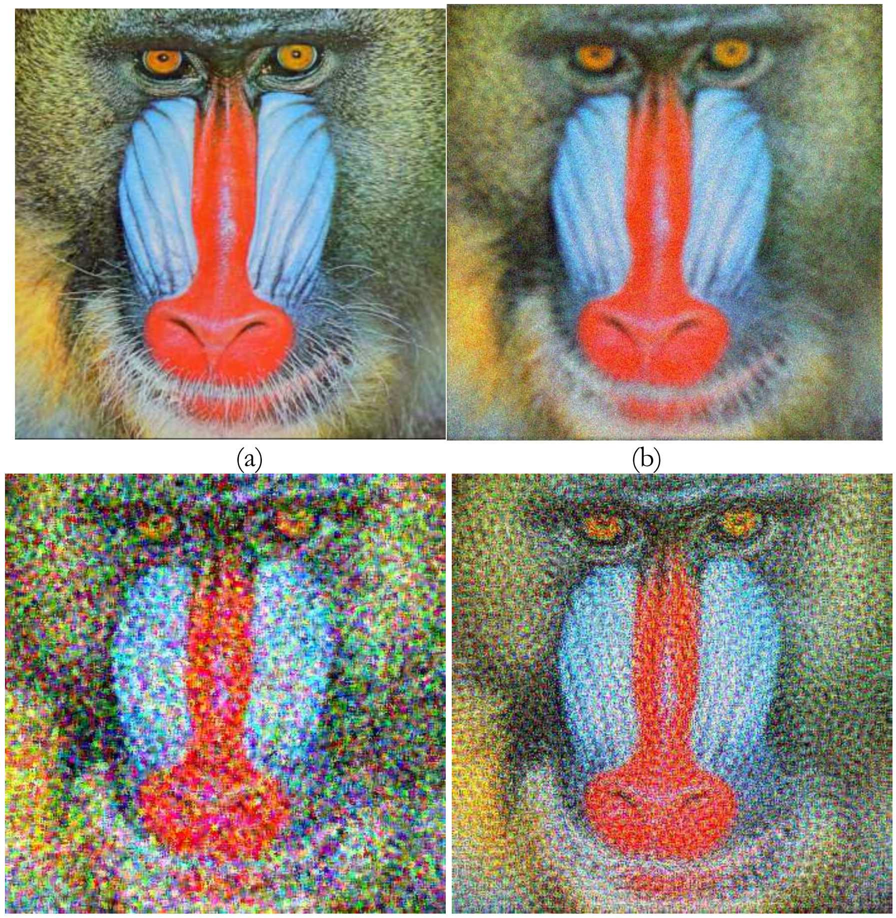


Figure 4.4: Interaction of denoise and deblur filters on a color image
(a) Original [73], (b) Image a with synthetic 10 pixel defocus blur and noise
(c) Image b deblurred then denoised with a median filter,
(d) Image b denoised with a median filter denoise with 2x2 pixel neighborhood and then deblurred

10) Deblur

If an image is affected by a blur, it is best to eliminate it before any sharpening or color corrections are made. Sharpening is an operation used to heighten high frequency image information, resulting in a clearer distinction between details. Because a blurred image is composed of low frequency information, sharpening will not be effective. Additionally, attempting to sharpen a heavily blurred image will not allow one to see the effects of the sharpening procedure and make it difficult to estimate appropriate settings. This may result in an appearance of over or under sharpening after the image is deblurred. Likewise, proper color adjustments cannot be accurately estimated while an image is blurred since color information diffused throughout the image.

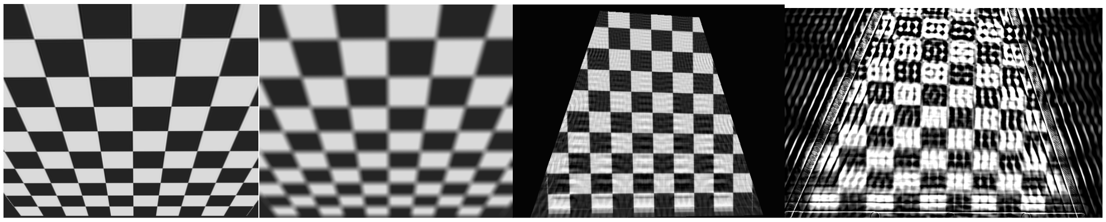
(a)

(b)

Figure: 4.5: Interaction of blur and sharpening
From left to right:
The first image [73] in figure (a) was blurred with a 25 pixel synthetic lens blur. This was then deblurred and sharpened. Image b was blurred with a 25 pixel synthetic lens blur. This was then "sharpened" and then deblurred. Though the resulting images are visually similar, this is only because both images where sharpened at the same settings. In practice it would be impossible to determine the correct amount of sharpening needed by visual analysis of the original blurred image.

11) Distortion Correction

Distortion correction changes the relationships between pixels and their neighbors, causing deblurring to be impossible afterwards. Deblurring algorithms typically work on the assumption that an image is uniformly blurred in one direction or from a central point. Distortion correction may cause areas of an image to be more blurred than others, impairing the efficiency of the deblurring algorithm. For these reasons, distortion correction should only be attempted after an image has been deblurred (if necessary).



(a) (b) (c) (d)

Figure 4.6: Interaction of distortion correction and deblur
(a) Original Image, (b) Original blurred with 20 pixel synthetic lens blur
(c) Image b deblurred then rectified, (d) Image b rectified then deblurred

(a) (b) (c) (d)

Figure 4.7: Interaction of distortion correction and deblur on a color image
(a) Original Image, (partial license plate obscured),
(b) Original blurred with 35 pixel synthetic motion blur,
(c) Image b deblurred then rectified, (d) Image b rectified then deblurred

12) Contrast/Color Correction

Color and contrast adjustments are more efficient if they are made after deblurring before to ensure that color is accurately represented and not convolved throughout the image. Furthermore, by adjusting the contrast between a group of pixels, details can be revealed which make it easier to identify areas for sharpening. Testing showed that color adjustments before and after deblurring had little impact on the output image (Fig. 4.8). However, de-noising should be performed first if the image requires deblur before color adjustments. For this reason, contrast and color correction is listed after denoise and deblur, and before sharpening in this order of operations.

Figure 4.8: Interaction of noise and histogram EQ

The input image for this test [73] is 50 frames of the same image with random noise at each frame. The following descriptions read left to right:

(a)Noisy video with random noise at each frame -> *a* with frame averaging -> averaged image with histogram EQ. (b) Noisy video with random noise at each frame -> a with histogram EQ -> Equalized image with frame averaging. (c)Noisy video with random noise at each frame -> *a* with median filter -> median filtered with histogram EQ. (d) Noisy video with random noise at each frame -> a with histogram EQ -> Equalized image with median filter.

13) Sharpening

If the previous steps were followed up to this point, noise should be diminished and details should be revealed with deblur and tonal adjustments, allowing for sharpening to be effective. Denoising does carry the possibility of removing actual image texture however, so an examiner may deem it necessary to forgo denoising and sharpen an image directly.

(a) (b) (c)

Figure 4.9: Interaction of sharpening and denoising
(a) A noisy image [73], (b) Image a sharpened then denoised with a median filter,
(c) Image a, first denoised with a median filter then sharpened.

<u>Output</u>

14) After enhancement, the image file must be output. Image should be exported with a non-proprietary, lossless encoding in order to ensure that image details are not lost and that the image or video that can be viewed easily. In addition to outputting the newly enhanced image data, lab best practices should be upheld to ensure provenance of the enhanced media, as is done during acquisition.

15) Though documentation should be kept at every single step of this framework, this is listed as the final step in order to ensure that steps which have not previously been tracked before are written down. This may include information not recorded in a software history log such as file output settings, file output hash values, and chain of custody information.

CHAPTER V

CONCLUSION

Through research conducted in this paper, it was found that the order in which certain enhancements operations are applied can affect the visual fidelity of a processed image. By conducting testing of various image enactment filter combinations, the following were determined: Interlaced video cannot be accurately de-interlaced after processing such as interpolation, blur, and distortion correction. Noise can negatively affect deblurring algorithms, contrast adjustments, and sharpening filters. Blur impacts the visibility of sharpening and color corrections. Deblur cannot be reliably accomplished after distortion corrections. Lastly, deblurring and adjusting color can help identify regions that can be improved by sharpening. This research was used to develop an order of operations for forensic image processing that maximizes the efficacy of each enhancement.

The proposed order of operations for forensic image enhancement presented in this paper attempts to find an ideal balance between increased image fidelity and artifact reduction by investigating the properties of images and interactions of multiple enhancement techniques.

Naturally, this order may not be ideal depending on the content of an image and the needs of a forensic investigation. In such cases, it may be necessary to deviate from this proposed framework. However in order to deviate responsibly, it is necessary to have a strong understanding of digital images, the underlying processes of image enhancement techniques, forensic principles, and a scientific perspective with which to validate novel methods. The first three chapters of this paper attempt to help the reader develop understanding in these topics so that he or she can understand the rationale behind the proposed framework for image enhancement presented in Chapter 4 and deviate from it in a forensically sound manner if necessary.

Granting that a fully ideal order for image processing will likely never be fully realized, what is important is the search for such an ideal. Image capture technology and image processing techniques will surely be unrecognizable in a short span of time, at which point much of the content of this paper may be obsolete. What will never be obsolete however is the ethical responsibility for forensic scientists to be mindful of human fallibility and to find new ways to mitigate error and yield reliable results. This work hopes to be a small step in that direction.

REFERENCES

[1] Scientific Working Groups on Imaging Technology, Section 11 Best Practices- for Documenting Image Enhancement. Available: https://www.swgit.org/pdf/Section%2011%20Best%20Practices%20for%20Documenting%20Image%20Enhancement?docID=37

[2] Jain, A. K. (1989). Fundamentals of digital image processing. Prentice-Hall, Inc..

[3] Forecast of digital cameras penetration in U.S. household from 2010 to 2013, Available: http://www.statista.com/statistics/198200/household-penetration-of-digital-cameras-in-the-united-states-since-2010/

[4] Schweitzer, N. J., & Saks, M. J. (2007). The CSI effect: Popular fiction about forensic science affects the public's expectations about real forensic science. Jurimetrics, 357-364.

[5] Cole, Simon; Dioso, Rachel (13 May 2005). Law and the Lab". The Wall Street Journal.

[6] Mirsky, S. (2005). Crime scene instigation. Scientific American, 292(5), 100-100.

[7] Shelton, Donald E. "The'CSI Effect': Does It Really Exist?." National Institute of Justice Journal 259 (2008).

[8] Cheng R. (2015), Moore's Law is the reason your iPhone is so thin and cheap, cnet.com, Available: http://www.cnet.com/news/moores-law-is-the-reason-why-your-iphone-is-so-thin-and-cheap/

[9] Carl Sagan’s last interview with Charlie Rose (Full Interview), Available: https://www.youtube.com/watch?v=U8HEwO-2L4w

[10] Sagan, C. (2011). Demon-haunted world: science as a candle in the dark. Ballantine Books.

[11] Erzinclioglu, Z. (2000). Every Contact Leaves a Trace: Scientific Detection in the Twentieth Century. Carlton Books Limited.

[12] Scientific Working Groups on Digital Evidence and Imaging Technology, SWGDE/SWGIT Digital & Multimedia Evidence Glossary. Available: https://www.swgit.org/pdf/SWGDE%20and%20SWGIT%20Digital%20and%20Multimedia%20Evidence%20Glossary?docID=60

[13] Anderson, S. D. (2011). Digital image analysis: Analytical framework for authenticating digital images (Graudate Thesis, University of Colorado Denver).

[14] BBC, Iran 'faked missile test image', Available: http://news.bbc.co.uk/2/hi/middle_east/7500917.stm

[15] Grigoras C., and Smith J.M. In: Siegel JA and Saukko PJ (eds.) Encyclopedia of Forensic Sciences, Second Edition. pp. 303-314. Digital Imaging: Enhancement and Authentication, Waltham: Academic Press.

[16] Scheinfeld, N. (2004). Photographic images, digital imaging, dermatology, and the law. Archives of dermatology, 140(4), 473-476.

[17] Jones P. In: Siegel JA and Saukko PJ (eds.) Encyclopedia of Forensic Sciences Second Edition, Pages 335–341 Digital Evidence/Photography and Digital Imaging, Waltham: Academic Press.

[18] Blitzer, H. L., Stein-Ferguson, K., & Huang, J. (2010). Understanding forensic digital imaging. Academic Press.

[19] Daniel, M. (2004). Daguerre (1787–1851) and the Invention of Photography.Heilbrunn Timeline of Art History.

[20] Federal Rules of Evidence

[21] Russ, J. C. (2001). Forensic uses of digital imaging. CRC Press.

[22] FindLaw, BOGOSIAN v. MERCEDES BENZ OF NORTH AMERICA INC, Available: http://caselaw.findlaw.com/us-1st-circuit/1372030.html

[23] OpenJurist, Frank LOCASCIO, and John Gotti, Defendants-Appellants. (1993) Available: http://openjurist.org/6/f3d/924/united-states-v-locascio

[24] Bronstein, D. A. (2011). Law for the expert witness. CRC Press.

[25] Smith, F. C., & Bace, R. G. (2002). A guide to forensic testimony: The art and practice of presenting testimony as an expert technical witness. Pearson Education.

[26] Kaufman, M. S. (2006). The status of Daubert in state courts. Atlantic Legal Foundation, (March 31).

[27] The Florida Senate Committee on Judiciary 2013 Summary of Legislation Past, HB 7015 – Expert Testimony. Available: http://www.flsenate.gov/Committees/BillSummaries/2013/html/489

[28] 2014 KBA Annual Meeting, Daubert Comes to Kansas: How the Federal Standard Will be Applied in State Court. Available: http://c.ymcdn.com/sites/www.ksbar.org/resource/resmgr/am14_materials/%287%29daubert.pdf

[29] Li, F. (2008). Probability of false positive with an innocent image processing routine. Journal of Forensic Identification, 58(5), 551.

[30] SWGIT Section 17. Digital Imaging Technology Issues for the Courts. Version 2.2 2012. Available:https://www.swgit.org/pdf/Section%2017%20Digital%20Imaging%20Technology%20Issues%20for%20the%20Courts?docID=56

[31] Nagosky, D. P. (2005). The admissibility of digital photographs in criminal cases. FBI Law Enforcement Bulletin, 74(12), 1.

[32] Robinson, E. M. (2009). Crime scene photography. Academic Press.

[33] FindLaw, RODD v. RARITAN RADIOLOGIC ASSOCIATES, Available: http://caselaw.findlaw.com/nj-superior-court-appellate-division/1446078.html%29%20%20T

[34] Levy-Sachs R., Sullivan R., Using Digital Photographs in the Courtroom - Considerations for Admissibility . Available: http://www.forensictv.net/Downloads/legal/using_digital_photographs_in_the_courtroom_considerations_for_admissibility_by_rebecca_levy_sachs_melissa_sullivan_robinson_and_cole.pdf

[35] Tilstone, W. In: Mozayani, A., & Noziglia, C. (Eds.). (2010). The forensic laboratory handbook procedures and practice. Springer Science & Business Media. pp. 335-367 Quality in the Forensic Science Laboratory.

[36] SWGIT Section Best Practices for Maintaining the Integrity of Digital Images and Digital Video. Version 1.1 Available: https://www.swgit.org/pdf/Section%2013%20Best%20Practices%20for%20Maintaining%20the%20Integrity%20of%20Digital%20Images%20and%20Digital%20Video?docID=54

[37] SWGDE, Minimum Requirements for Quality Assurance in the Processing of Digital and Multimedia Evidence. Available: https://www.swgde.org/documents/Current%20Documents/2010-05-15%20SWGDE%20Min%20Req%20for%20QA%20in%20Proc%20Digital%20Multimedia%20Evidence_v1

[38] Smith, J. Quality Assurance in DME. Lecture. National Center for Media Forensics

[39] Edwards, H., & Gotsonis, C. (2009). Strengthening forensic science in the United States: A path forward. *Statement before the United State Senate Committee on the Judiciary*.

[40] SWGDE Scientific Working Group on Digital Evidence (SWGDE) Position on the National Research Council Report to Congress Strengthening Forensic Science in the United States: A Path Forward . Available: https://www.swgde.org/documents/Current%20Documents/2010-05-15%20SWGDE%20Min%20Req%20for%20QA%20in%20Proc%20Digital%20Multimedia%20Evid Evid_v1

[41] National Institute of Science and Technology, (2014), Summary of the NIST Proposed Plan for the Organization of Scientific Area Committees. Available: http://www.nist.gov/forensics/upload/OSACsummary3-18-14.pdf

[42] National Institute of Science and Technology, (2014), NIST Proposed Organization of Scientific Area Committees (OSAC). Available: http://www.nist.gov/forensics/upload/NIST-OSAC-Plan-NCFS-Feb-4-2014-2-3-14-FINAL.pdf

[43] Damjanovski, V. (2013). CCTV: From Light to Pixels. Elsevier.

[44] Wyszecki, G., & Stiles, W. S. (1982). Color science (Vol. 8). New York: Wiley.

[45] Gregory, R., & Cavanagh, P. (2011). The blind spot. Scholarpedia, 6(10), 9618.

[46] Gehring, W. J. (2004). Historical perspective on the development and evolution of eyes and photoreceptors. Int J Dev Biol, 48(8-9), 707-17.

[47] Williamson, S. J., & Cummins, H. Z. (1983). Light and color in nature and art.Light and Color in Nature and Art, by Samuel J. Williamson, Herman Z. Cummins, pp. 512. ISBN 0-471-08374-7. Wiley-VCH, February 1983., 1.

[48] Kerr, R., Needham, J., & Wood, N. (2004). Science and Civilisation in China: Volume 5, Chemistry and Chemical Technology, Part 12, Ceramic Technology. Cambridge University Press.

[49] Kansal, N. K. (2010). Fuzzy techniques for image enhancement (Doctoral dissertation, Thapar University Patiala)

[50] Guttosch, R. J. (2002). Investigation of Color Aliasing of High Spatial Frequencies and Edges for Bayer-Pattern Sensors and Foveon X3® Direct Image Sensors. Foveon Inc. http://www. foveon. com/files/Color_Alias_White_Paper_FinalHiRes. pdf.

[51] Snavely N., CS6670 Computer Vision. Available: http://www.cs.cornell.edu/courses/cs6670/2011sp/lectures/lec02_filter.pdf

[52] Maini, R., & Aggarwal, H. (2010). A comprehensive review of image enhancement techniques. arXiv preprint arXiv:1003.4053.

[53] Judd, D. B., & Wyszecki, G. (1963). Color in business, science and industry.

[54] Cambridge In Colour, Bit Depth Tutorial. Avaialble; http://www.cambridgeincolour.com/tutorials/bit-depth.htm

[55]Haberdar, H. "Spatial Frequency of an Image Tutorial", Available: http://www.haberdar.org/Spatial-Frequency-of-an-Image-Tutorial.htm

[56] EC Council, (2010) Computer Forensics: Investigating Data and Image Files, 1st Edition

[57] Setton, E., & Girod, B. (2005, July). Video streaming with SP and SI frames. InVisual Communications and Image Processing 2005 (pp. 59606F-59606F). International Society for Optics and Photonics.

[58] Lancaster, D. (2007). A review of some image pixel interpolation algorithms.copyright c2007 as Guru Gram, 32.

[59] Cambridge in Colour, Digital Image Interpolation. Available: http://www.cambridgeincolour.com/tutorials/image-interpolation.htm

[60] BBC News, ‘Swirl’ suspect trial adjouned.(2008), Available: http://news.bbc.co.uk/2/hi/asia-pacific/7286815.stm

[61] ColorPicker, Available: http://www.colorpicker.com/

[62] Bedi, S. S., & Khandelwal, R. (2013). Various image enhancement techniques-a critical review. International Journal of Advanced Research in Computer and Communication Engineering, 2(3).

[63] Cambridge in Colour, digital camera image noise. Available: http://www.cambridgeincolour.com/tutorials/image-noise.htm

[64] Ludwig, J. (n.d.), "Image Convolution". Portland State University.http://web.pdx.edu/~jduh/courses/Archive/geog481w07/Students/Ludwig_ImageConvolution.pdf

[65] Color Cambridge in Colour, Guide to Image Sharpening. Available: http://www.cambridgeincolour.com/tutorials/image-sharpening.htm

[66] MatlabTricks, 3x3 convolution kernels with online demo (2013). Available: http://matlabtricks.com/post-5/3x3-convolution-kernels-with-online-demo#demo

[67] MathWorks, Understanding Deblurring, http://www.mathworks.com/help/images/understanding-deblurring.html

[68] Al-Ameen, Z., Sulong, G., & Johar, M. G. M. Toward a Better Understanding of Deblurring Digital Images in Computer Forensics.

[69] Park, S. C., Park, M. K., & Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview. Signal Processing Magazine, IEEE, 20(3), 21-36.

[70] Smith, J. Image Enhancemnet. Lecture. National Center for Media Forensics

[71] Reis, G. (2007). Photoshop CS3 for Forensics Professionals: A Complete Digital Imaging Course for Investigators. John Wiley & Sons.

[72] Al-Ameen, Z., Mohamad, D., & Sulong, G. (2012). Restoring degraded astronomy images using a combination of denoising and deblurring techniques. *International Journal of Signal Processing, Image Processing and Pattern Recognition (IJSIP)*,*5*(1).

[73] USC, The USC-SIPI Image Database, Available: http://sipi.usc.edu/database/